



António Duarte Zapico Bicho de Sousa Franco

Licenciado em Bioquímica

Long non-coding RNA contributes to direct conversion reprogramming

Dissertação para obtenção do Grau de Mestre em
Genética molecular e biomedicina

Orientador: Bruno Bernardes de Jesus, PhD, IMM
Co-orientador: Maria do Carmo Fonseca, PhD, IMM/FM-UL

Júri:

Presidente: Prof. Doutor José Paulo Nunes de Sousa Sampaio
Arguente: Doutora Manuela Cristina Fernandes Ferreira
Vogal: Doutor Bruno Miguel Bernardes de Jesus



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2016



Long non-coding RNA contributes to direct conversion reprogramming
António de Sousa Franco



António Duarte Zapico Bicho de Sousa Franco

Licenciado em Bioquímica

Long non-coding RNA contribute to direct conversion reprogramming

Dissertação para obtenção do Grau de Mestre em
Genética molecular e biomedicina

Orientador: Bruno Bernardes de Jesus, PhD, IMM
Co-orientador: Maria do Carmo Fonseca, PhD, IMM/FM-UL

Júri:

Presidente: Prof. Doutor José Paulo Nunes de Sousa Sampaio
Arguente: Doutora Manuela Cristina Fernandes Ferreira
Vogal: Doutor Bruno Miguel Bernardes de Jesus



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2016

Long non-coding RNA contribute to direct conversion reprogramming

Copyright António Duarte Zapico Bicho de Sousa Franco, FCT/UNL, UNL

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Agradecimentos

Quero expressar o meu agradecimento ao meu orientador Bruno Bernardes de Jesus pela oportunidade de participar neste projeto assim como por me ter acolhido e guiado ao longo desta tese, sem o qual o seu desenvolvimento teria sido impossível, e também à minha coorientadora Maria do Carmo Fonseca, professora da FM-UL (Faculdade de Medicina da Universidade de Lisboa) e chefe do laboratório de *RNA & Gene Regulation* do IMM (Instituto de Medicina Molecular), por me ter aceite no seu laboratório e permitido desenvolver este trabalho.

Quero também agradecer aos técnicos da FM-UL Sérgio Pires Marinho e José António Albuquerque pelo constante apoio e aconselhamento que me deram no início desta tese e ao longo da mesma, no planeamento e execução das técnicas laboratoriais elaboradas, assim como aos restantes colegas e amigos no IMM que me acompanharam ao longo deste ano e que de uma forma ou de outra contribuíram para a elaboração do trabalho adiante apresentado.

Por ultimo mas não menos importante, quero agradecer à minha família e amigos pelo apoio prestado, principalmente aos meus pais por sempre acreditarem nas minhas capacidades.

A todos os já referidos e mais alguns que por esquecimento meu possam não ter sido, deixo o meu muito obrigado.

Resumo

O aumento do conhecimento em anos recentes levou os RNAs longos não codificantes a saltar de ruído para importantes reguladores da expressão genética, sabendo-se agora do seu envolvimento em várias vias de regulação onde atuam de diversas formas. Alguns destes RNAs são importantes em processos de diferenciação celular, pela sua expressão diferencial e temporal nos diversos tecidos e fases evolutivas do organismo. A descoberta de que células terminalmente diferenciadas podem reverter a estados de multipotência e pluripotência através da expressão forçada de genes específicos abriu a porta ao mundo da reprogramação celular. A participação de vários RNAs longos não codificantes em processos de regulação sugere que estas moléculas possam participar ativamente em reprogramação pelo que identificar os alvos corretos e como manipulá-los pode atenuar algumas das limitações de alguns protocolos existentes. RNAs longos não codificantes com influência em processos de transição como EMT ou MET podem ser importantes em reprogramação direta, e a expressão de RNAs não codificantes característicos de determinada linhagem celular podem também contribuir. Neste trabalho são explorados dois RNAs não codificantes envolvidos neste tipo de processos, Zeb2NAT e Pnky, e a sua potencial aplicação na reprogramação direta de fibroblastos em células multipotentes pela expressão de apenas um fator de transcrição. Limitar a expressão do Zeb2NAT durante os primeiros passos de reprogramação parece atenuar o processo, no entanto se aumentar a sua expressão poderá ter um efeito inverso está ainda por determinar. A obtenção de células progenitoras de diferentes linhagens a partir de células terminais do próprio paciente apresenta potenciais médicos enormes, e a utilização de RNAs longos não codificantes em reprogramação celular pode ser um valioso auxílio ou alternativa na procura de contornar as limitações existentes em reprogramação ou a dificuldade acrescida em reprogramar células de pacientes de idade avançada.

Palavras-chave: RNAs longos não codificantes; Reprogramação celular; Multipotência; Zeb2NAT; Pnky

Abstract

Over the last years, research led lncRNAs to go from transcriptional noise to important regulators of gene expression, being now known their association with many regulatory pathways by a wide set of mechanisms. Some of these RNAs play very important roles in cell differentiation, given their differential and temporal expression during tissue and organismal development. The discovery that terminally differentiated cells could revert to pluripotency or be redirected to multipotent progenitors of different lineages through forced expression of a specific onset of genes opened the way for direct conversion studies. The fact that various lncRNAs have been associated with differentiation, pluripotency and cancer indicates that these molecules might be useful tools in direct conversion. Manipulation of lncRNA expression during cell reprogramming could provide aid in overcoming current protocol limitations as the tumorigenic potential of iPSCs, low efficiencies and aging related epigenetic resistance. LncRNAs with influence in cell character transitions such as epithelial-mesenchymal transition (EMT) or mesenchymal-epithelial transition (MET) could contribute to direct conversion reprogramming, as could lncRNAs that are specifically expressed in the reprogramming target cell line. We herein present two lncRNAs with the mentioned characteristics, Zeb2NAT and Pnky, respectively, as potential targets for improving fibroblast direct conversion reprogramming to multipotent hematopoietic and neural progenitors with a single pluripotency transcription factor (TF). Downregulation of Zeb2Nat during direct conversion seems to impair reprogramming efficiency, however if upregulating this lncRNA will have an improving effect is still under study. Getting unlimited patient-specific progenitor cells of different lineages from more accessible sources presents an enormous medical issue.

Keywords: LncRNA; direct reprogramming; multipotency; Zeb2NAT; Pnky

Contents

Agradecimientos.....	I
Resumo.....	III
Abstract.....	V
List of abbreviations, acronyms and symbols.....	XIII
1. Introduction.....	1
I. Induced pluripotency, the beginning of reprogramming.....	1
II. Direct conversion and trans-differentiation.....	2
III. Direct conversion as a medical alternative to iPSCs.....	3
IV. Non-coding RNAs and their role in lineage commitment.....	4
V. LncRNA regulation as a potential direct reprogramming enhancer.....	6
2. Materials and Methods.....	11
Reprogramming protocols.....	11
Retroviral production.....	11
Zeb2-NAT downregulation.....	11
Culture Conditions.....	12
a) Fibroblast culture.....	12
b) iHPC.....	12
c) iNSC.....	12
Reprogrammed cells analysis.....	12
a) iHPC.....	12
b) iNSC.....	12
Pnky cloning.....	13
Pnky retroviral plasmid construction.....	13
a) pMXs Backbone (sticky ends).....	13
b) pBabe Backbone (blunt ends ligation).....	14
3. Results and discussion.....	15
Zeb2NAT in early reprogramming.....	15
iHPC.....	16
iNSC.....	20
LncRNA overexpression by retroviral expression.....	22
Future Prospects.....	24
4. Conclusions.....	25
5. References.....	27
6. Annexes.....	i
Primer Sequences (5'-3').....	i
LNA GapmeR sequences (5'-3').....	i
mPnky cDNA sequence.....	ii
iHPCs.....	iii
iNSCs.....	iv

Index of Figures

Figure 1.1 - Schematic representation of iPSC characterization (De Los Angeles et al., 2015).....	2
Figure 1.2 - Fibroblast reprogramming by ectopic expression of pluripotency TFs. Transient expression of iPSC TFs leads to the appearance of epigenetically “activated” cells capable of expressing lineage specific genes under medium stimulation, allowing direct conversion to lineage committed precursors that can be further differentiated to terminal cell lines for biomedical application (<i>Tianhua Ma et al. 2013</i>).....	3
Figure 1.3 - Non-coding RNA regulation of mesenchymal-epithelial transition by controlling E-Cadherin expression. Zeb2NAT is induced by Snail1 promoting ZEB2 processing which will repress E-Cadherin leading to EMT (<i>Pauli et al, 2011</i>).....	5
Figure 1.4 - Illustration of LNA GapmeR™ action (Exiqon.com)	6
Figure 1.5 - A Representation of Pnky localization and influence in neurogenesis from NSCs (<i>Ramos et al 2015</i>); B Schematic representation of neuronal fate commitment regulation by PTB proteins. REST and PTBP1 repress neuronal differentiation of neural stem/progenitor cells, while PTBP2 promotes neuron maturation (<i>adapted from Kawahara et al 2012</i>).....	7
Figure 1.6 - A Representation of Zeb2 role in neuroectoderm formation by inducing neural genes and repressing epidermal and mesoendodermal fate B Schematic representation of Zeb2 knockout in cranial, vagal and trunk neural crest cell populations effect on NCCs formation, migration and specification (<i>Hegarty et al, 2015</i>)	8
Figure 3.1 - Downregulation efficiency assay, Zeb2NAT expression level in HuFs treated with LNA GapmeRs targeting Zeb2NAT transcripts (α NAT) compared to wild type HuFs (Control).....	15
Figure 3.2 - Human fibroblast cell lines WI-38(A), 3YR (B), 11YR (C) and 73YR (D) transduced with GFP retrovirus for transduction efficiency positive control (Photos were taken in Zeiss Axiovert 200M microscope with x10 amplification).....	16
Figure 3.3 - Experiment model for fibroblast reprogramming with Oct4 (Oct4 - HuFs transduced with hOct4; LNA - HuFs treated with scramble LNA; α NAT - HuFs treated with LNA GapmeRs targeting Zeb2NAT; GFP - HuFs transduced with GFP; WT - Wild type HuFs). Conditions in black are for HuFs cultured in differentiation medium and in brown for fibroblast culture conditions during 21 days after retroviral transduction and LNA treatment.	17
Figure 3.4 - WI-38 human fetal fibroblasts transduced with Oct4 alone (A), with Oct4 and treated with Zeb2NAT LNA GapmeRs (B) or scramble LNAs (C) at day 5 (1), day 10 (2) and day 19 (3) in differentiation medium composed of KSR medium supplemented with IGF2 and FGF2. WI-38 cells transduced with GFP alone as control at day 5 (D1) and day 19 (D2) in the same medium (photos taken with Zeiss PrimoVert Microscope, amplification x4 (1) and x10 (2, 3)).	18
Figure 3.5 - FACS profile controls. CD45 positive control cells unstained (A1) and stained with anti-humanCD45 antibody (A2), HuFs transduced with Oct4 unstained (B) and WT HuFs stained for CD45 (C) for negative control.....	18
Figure 3.6 - WI-38 HuFs FACS profile when cultured in KSR medium supplemented with IGF2 and FGF2 transduced with Oct4 alone (A,B and C) and treated with Zeb2NAT LNA GapmeRs (B) or scramble LNAs (C) and stained for CD45 (1-Unstained, 2-Stained for CD45). FACS analysis performed using BD Accuri C6 with a cell count of 10000.	19
Figure 3.7 - 3YR human fibroblasts transduced with Oct4 alone (A), with Oct4 and treated with Zeb2NAT LNA GapmeRs (B) or scramble LNAs (C) at day 6 (1), day 14 (2) and day 21 (3) in differentiation medium composed of KSR medium supplemented with IGF2 and FGF2. WI-38 cells transduced with GFP alone as control at day 6 (D1) and day 21 (D2) in the same medium (photos taken with Zeiss PrimoVert Microscope, amplification x4 (1) and x10 (2, 3)).	19
Figure 3.8 - 3YR HuFs cultured in KSR medium supplemented with IGF2 and FGF2 transduced with Oct4 alone (A,B and C) and treated with Zeb2NAT LNA GapmeRs (B) or scramble LNAs (C) and stained for CD45 and CD34. 3YR HuFs transduced with GFP retrovirus alone and cultured under the same conditions were also stained for negative control (D). Cell count=5000.....	20
Figure 3.9 - Experimental model for fibroblast direct reprogramming with Sox2. All HuFs in this experiment were cultured for 10 days in ReNcell medium supplemented with EGF and FGF2 at	

20ng.mL ⁻¹ , after retroviral transduction and LNA treatment (Sox2 - HuFs transduced with hSox2; LNA - HuFs treated with scramble LNA; α NAT - HuFs treated with Zeb2NAT targeting LNA; GFP - HuFs transduced with GFP).....	21
Figure 3.10 - A WI-38 HuFs transduced with Sox2 alone (A), with Sox2 and treated with LNA GapmeRs targeting Zeb2NAT (B) and with scramble LNAs (C) after 3 days (1), 5 days (2) and 9days (3) in culture with neural differentiation medium (for control cells see Figure D in supplementary data); B RT-qPCR result expression levels, presented as $\Delta\Delta C_t$ normalized using GFP retrovirus only transduced cells as control and GAPDH as housekeeping gene (for results using Actin as housekeeping see supplementary data) (photos taken with Zeiss PrimoVert Microscope, amplification x10).....	21
Figure 3.11 - A 73YR HuFs transduced with Sox2 alone (A), with Sox2 and treated with LNA GapmeRs targeting Zeb2NAT (B) and with scramble LNAs (C) after 3 days (1), 5 days (2) and 9days (3) in culture with neural differentiation medium (for control cells see Figure D in supplementary data); B RT-qPCR result expression levels, presented as $\Delta\Delta C_t$ normalized using GFP retrovirus only transduced cells as control and GAPDH as housekeeping gene (for results using Actin as housekeeping see supplementary data) (photos taken with Zeiss PrimoVert Microscope, amplification x10).....	22
Figure 3.12 - H3K4me3 and H3K27me3 ChIP-seq tracks for embryonic stem cells (ES), mouse embryonic fibroblasts (MEF), ESC-derived neural stem cells (ESC-NSCs), and V-SVZ NSCs of the Pnky and Pou3f2 loci viewed in Genome browser with promoter regions highlighted in yellow (Ramos et al. 2015).....	23
Figure 3.13 - A PCR product of SVZ cDNA amplified with mPnky homology primers and ligation ends with BstXI restriction sites, purified and run on agarose 1% TAE gel. B pBabeGFP plasmid design (addgene), plasmid was restricted with EcoRI and ClaI, excised GFP was substituted by mPnky.	23
Figure 6.1 - Colony forming from Oct4 transduced 3YR HuFs after 20 days in differentiation culture conditions and FACS analysis profile of same cells stained with anti-humanCD45 and anti-humanCD34 antibodies.	iii
Figure 6.2 - Histogram representation of 3YR cells reprogramed with Oct4 (Red for unstained and Blue for cells stained with APC-antiCD45 antibody). A Oct4 alone, B Oct4 + LNA targeting Zeb2NAT and C Oct4 + Scramble LNA.....	iii
Figure 6.3 - Control WI-38 HuFs transduced with GFP retrovirus alone after 3 days (D1) and 9days (D2) in culture with neural differentiation medium and RT-qPCR result expression levels, presented as $\Delta\Delta C_t$ using Actin as housekeeping gene.	iv
Figure 6.4 - Control 73YR HuFs transduced with GFP retrovirus alone after 3 days (D1) and 9days (D2) in culture with neural differentiation medium and RT-qPCR result expression levels, presented as $\Delta\Delta C_t$ using Actin as housekeeping gene.....	iv

Index of Tables

Table 2.1 - Brief description of human cell lines used for reprogramming experiments.....	11
---	----

List of abbreviations, acronyms and symbols

LncRNA – Long noncoding RNA
miRNA – Micro RNA
LincRNA – Long intergenic noncoding RNA
MET – Mesenchymal to epithelial transition
EMT – Epithelial to mesenchymal transition
ICM – Inner cell mass
ESC – Embryonic stem cell
hESC – Human embryonic stem cell
mESC – Mouse embryonic stem cell
iPSC – Induced pluripotent stem cell
hiPSC – Human induced pluripotent stem cell
miPSC – Mouse embryonic stem cell
MEF – Mouse embryonic fibroblast
TF – Transcription factor
mEpiSC – Murine epiblast stem cell
FGF – Fibroblast growth factor
TGF β – Transforming growth factor β
LIF – Leukemia inhibitory factor
Mek – Mitogen-activated protein kinase (*aka MAPK*)
Gsk3 – Glycogen synthase kinase 3
NSC – Neural stem cell
iNSC – Induced neural stem cell
HPC – Hematopoietic progenitor cell
HSC – Hematopoietic stem cell
iHPC – Induced hematopoietic progenitor cell
iHSC – Induced hematopoietic stem cell
EPC – Endothelial progenitor cell
SVZ – Sub-ventricular zone
V-SVZ – Ventricular sub-ventricular zone
IRES – Internal ribosome entry site
NAT – Natural antisense transcript
IGF2 – Insulin-like growth factor 2
FLT3 – Fms-like tyrosine kinase 3
FLT-LG – Fms-like tyrosine kinase 3 ligand
SCF – Stem cell factor
IL – Interleukin
BMP – Bone morphogenetic protein
G-SCF – Granulocyte-colony stimulating factor
EGF – Epidermal growth factor
CD – Cluster of differentiation
Sip – Smad interacting protein
CtBP – C-terminal Binding Protein
NuRD – Nucleosome Remodeling Deacetylase
Nestin – Neuroectodermal stem cell marker
Tuj1 – Neuron-specific class III beta-tubulin
MAP2 – Microtubule-associated protein 2
PTB – Polypyrimidine tract-binding
PTBP - Polypyrimidine tract-binding protein
REST – RE1-Silencing transcription factor
GFAP – Glial fibrillary acidic protein
Nurr1 – Nuclear receptor related 1 protein

GFP – Green fluorescent protein
DMEM – Dulbecco's Modified Eagle Medium
KSR – Knock-out serum replacement
FBS – Fetal bovine serum
PBS – Phosphate-buffered saline
LB – Lysogeny broth
NEAA – Non-essential amino acids
Amp – Ampicillin
Pen – Penicillin
Strep – Streptomycin
LNA – Locked nucleic acids
FACS – Fluorescence-activated cell sorting
EDTA - Ethylenediaminetetraacetic acid
TAE – Tris-Acetic acid-EDTA
UTR – Untranslated region
ORF – Open reading frame

1. Introduction

I. Induced pluripotency, the beginning of reprogramming

Embryonic stem cells, originally derived from the inner cell mass (ICM) of the mammalian embryonic blastocyst, possess a vast potential in biomedical applications due to their capacity to differentiate into any of the 3 germ layer cells. These applications include tissue recovery and patient specific disease models (as in degenerative diseases), tissue specific toxicology assays, tissue *in vitro* production or personalized medicine. However, ethical issues in the use of human embryos as sources of stem cells constitute a barrier to their clinical application. Therefore, the possibility of devolving somatic cells to a pluripotent state, permitting the culture of cells with similar characteristics to embryonic stem cells (ESC) was a revolutionary discovery with a huge potential in the field of biomedicine (Takahashi & Yamanaka 2006; Yamanaka et al. 2007).

The possibility of obtaining induced pluripotent stem (iPS) cells was first demonstrated by *Yamanaka et al* via retroviral expression in murine embryonic fibroblasts (MEFs) of a combination of four transcription factors (TFs), Oct4 (*octamer-binding protein 4*), Sox2 (*SRY-box containing gene 2*), Klf4 (*Kruppel-like factor 4*) and c-Myc (*c-Myelocytomatosis oncogene*) (OSKM). Yamanaka and colleagues obtained cell colonies that were indistinguishable from ESCs in morphology, proliferation and gene expression (Takahashi & Yamanaka 2006). The first human induced pluripotent stem cells (hiPSC) were produced in laboratory, also by *Yamanaka et al*, using the same combination of transcription factors OSKM in adult human dermal fibroblasts. It is now known that these transcription factors originally used do not share an identical contribution to the reprogramming process of inducing pluripotency. The Oct4 and Sox2 genes present crucial functions being associated to the regulation and maintenance of the undifferentiated state, whereas Klf4 is involved in alterations of the chromatin structure by interacting, for example, with the histone acetyltransferase p300 enzyme, so that OCT4 and SOX2 can reach their genomic targets. On the other hand, cMyc is a well-known oncogene that contributes to iPSC proliferation in culture but has been shown to have a lesser contribution to the reprogramming process itself, having been recorded to sometimes even show adverse effects on the reprogrammed cells (Yamanaka et al. 2007).

To confirm the pluripotent capacity of reprogrammed cells, some conditions have to be met (see **Figure 1.1**). The first one is the expression of pluripotency markers like Oct4, Sox2, Nanog and SSEA (stage-specific embryonic antigen), another is the capacity of self-renewing accessed by cloning capacity in culture and embryoid body formation and, lastly is the ability to differentiate to all 3 germ lines accessed by teratoma formation when iPSCs are implanted *in vivo*, as do ESCs. Teratomas are tumor cell masses that rise from a disorganized expansion and differentiation process of pluripotent cells, producing the 3 germ layer cell lines, mesoderm, ectoderm and endoderm. The formation of this cell masses *in vivo* from iPSCs demonstrates their capacity to differentiate to the various cell lines (Abad et al. 2013).

There are two states of pluripotency in ESC and iPSCs according to slight differences in differentiation and proliferation character. These states are called the *naïve* and *primed* states, and are dependent on expression and epigenetic alterations during embryonic development that despite being common between mammals present variations in different species and embryo developmental states. The *naïve* state is a “ground state” characteristic of ESCs derived from the ICM of murine blastocysts, or murine embryonic stem cells (mESC), in which the general gene expression presents some differences to human embryonic stem cells (hESC), the latter showing a higher similarity to stem cells derived from the murine epiblast stage, differences such as the expression of additional transcription factors and non-methylated X chromosomes in female embryos. Since by comparison, hESCs seem to present an advanced developmental state and diminished pluripotency similar to the pluripotent capacity observed in murine epiblast stem (mEpiS) cells, they are considered *primed* stem cells, which is the characteristic state of mEpiS cells. *Primed* stem cells, as are hiPSCs, grow slowly in culture giving rise to flat colonies that depend on FGF2 (*fibroblast growth factor 2*) and TGFβ (*transforming growth factor β*) to self-renew. While *naïve* iPSCs depend on LIF (*Leukemia inhibitory factor*) and BMP4 (*Bone morphogenetic protein 4*) to grow in culture and form compact colonies that grow faster *in vitro* (Nichols & Smith 2009). However, hiPSCs may recover the *naïve* state if cultured with LIF when inhibitors of Mek and Gsk3 kinases are present in the medium (2i) (De Los Angeles et al. 2015). For biomedical interests, obtaining hiPSCs in the *naïve* state presents a greater medical application potential, therefore most studies sought to reprogram them to this state.

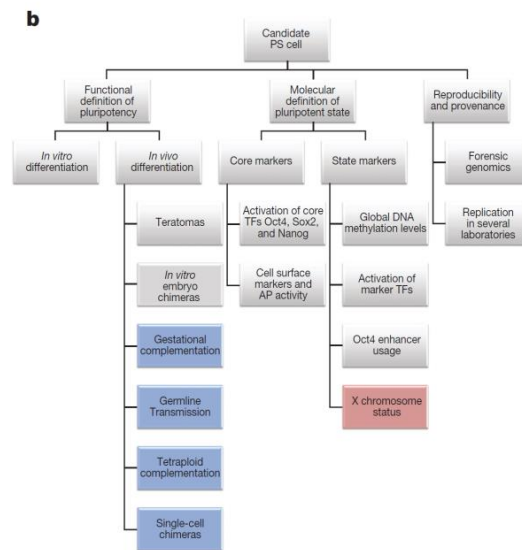


Figure 1.1 - Schematic representation of iPSC characterization (De Los Angeles et al. 2015)

II. Direct conversion and trans-differentiation

The discovery that ectopic expression of key transcription factors into terminally differentiated cells can revert them to an undifferentiated pluripotent state showed that terminally differentiated states were not as inflexible or irreversible as it was previously thought, raising questions on whether the possibility of directly turning one cell line into another would really be such a hard task to perform. This opened the door to direct reprogramming or direct conversion experiments based on a similar TF mediated approach. Direct conversion and trans-differentiation, although they are very similar in their purpose, present some important differences. Trans-differentiation is based on the over-expression of target cell specific TFs to direct cells towards a different cell line by activating a differentiation program, whereas direct conversion uses TFs related to pluripotency to promote the rise of epigenetically activated cells, inducing a partially undifferentiated state only to redirect them towards another cellular path by external stimulation before a fully pluripotent state is reached, skipping the intermediate iPSC state step. Also, trans-differentiation tends to produce terminal cell lines while direct conversion allows the appearance of primary lineage specified cells or progenitor cells, which present lineage-committed multipotent capacity and are able to maintain this state or differentiate even further into terminal and tissue-specific cells of that same lineage (Kelaini et al. 2014).

During recent years, various studies have confirmed the capacity of differentiated cells to directly change their cellular character without need to regress to an intermediate pluripotent state, thru a short period of OSKM expression before changing to conditions that favor differentiation to a specific

cell line of interest, for example, *Matsui et al* obtained iNSC from incompletely reprogrammed fibroblasts (Matsui et al. 2012). **Figure 1.2** presents a simplified model of direct conversion reprogramming based on this transient expression of OSKM (Ma et al. 2013).

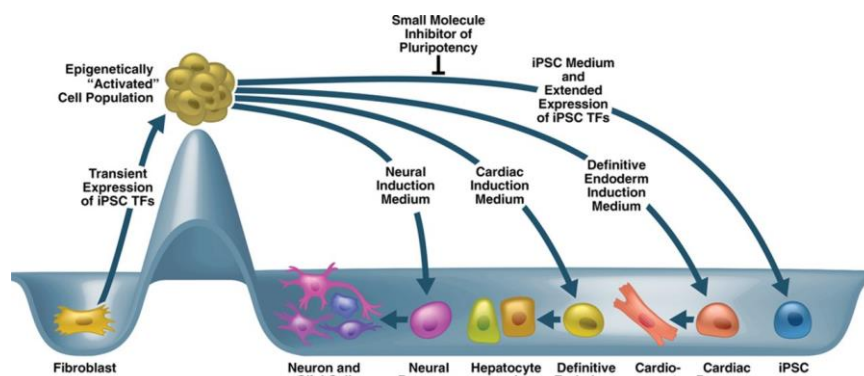


Figure 1.2 - Fibroblast reprogramming by ectopic expression of pluripotency TFs. Transient expression of iPSC TFs leads to the appearance of epigenetically “activated” cells capable of expressing lineage specific genes under medium stimulation, allowing direct conversion to lineage committed precursors that can be further differentiated to terminal cell lines for biomedical application (*Tianhua Ma et al. 2013*)

Inclusively, some studies have succeeded in performing direct conversion by ectopic expression of only one of the *Yamanaka* factors along with stimulation from medium supplementation (Szabo et al. 2010). Studies of this nature led to the attainment of induced multipotent cell lines derived from somatic cells (such as fibroblasts), as induced hematopoietic progenitor cells (iHPCs) (Szabo et al. 2010) or induced neural stem cells (iNSCs) (Li et al. 2013). Neural stem cells are glial cells that develop in the human embryonic central nervous system and are capable of self-renewing, long-term proliferation and differentiation into neurons, astrocytes and oligodendrocytes. In the adult brain, NSCs exist in two niches, one in the ventricular / sub-ventricular zone (V-SVZ) of the brain and other in the hippocampal dentate gyrus.

On the other hand, progenitor cells like HPCs are multipotent lineage-committed cells normally derived from hematopoietic stem cells (HSCs) of the bone marrow. Being one step ahead in differentiation progress, progenitor cells do not share the same self-renewing capacities of their committed stem cells of origin, although they are still not terminally differentiated, maintaining multipotent and proliferation capacities to some level.

III. Direct conversion as a medical alternative to iPSCs

There are still considerable concerns in the application of iPSC derived cells in clinical trials, some of them regarding possible resistance to differentiation due to incomplete or partial differentiation or epigenetic memory, which can lead to tumorigenesis in individuals implanted with iPSC derived cells. Another concern lies in the use of viral vectors to integrate transgenes during reprogramming and the risk of reactivation of their expression later on. To avoid this risk many studies search for alternative means to reprogram cell fate without relying on viral vectors (Ben-David & Benvenisty 2011).

Given today’s obstacles in applying the iPSC technology to medicine, in particular the possibility of tumor generation *in vivo*, obtaining multipotent cells by direct reprogramming skipping the pluripotent state is becoming a much more promising possibility on a short term basis, since induced

multipotent cells obtained from most current studies do not show apparent tumorigenic potential but maintain the same advantages of iPSC in what comes to personalized medicine applications.

As mentioned before, some studies have demonstrated to be possible to obtain multipotent or progenitor cells of specific lineages thru induced expression of only one key transcription factor in fibroblasts when in association with extracellular stimuli. Szabo *et al* discovered that human fibroblasts that only overexpress the Oct4 gene in iPSC culture conditions gave rise to colonies that expressed the pan-leucocyte marker CD45, and when supplementing the medium with growth factors IGF2, FGF2, SCF and the FLT3-ligand that stimulate the activation/appearance of hematopoietic CD (*clusters of differentiation*) surface proteins, these colonies grew in number. Then, exposing these CD45⁺ cells to an expansion medium supplemented with a cocktail of cytokines and growth factors (SCF, G-CSF, FLT3LG, IL-3, IL-6 and BMP-4) promoted the emergence of CD45 and CD34 co-expressing progenitor cells, capable of differentiating into macrophage, myeloid and megakaryocytic cells when further expanded in separate appropriate conditions, and were capable of *in vivo* engraftment. Although Oct4 has not been associated with hematopoietic fate, it shares similar binding motifs with Oct1 and Oct2, which have been previously implied in lymphoid development, making it possible for it to act by affinity on some hematopoietic related gene targets (Szabo *et al.* 2010).

Other studies also demonstrated the possibility to obtain multipotent capacity from fibroblasts by overexpression of a single pluripotency associated TF. Ring *et al* showed that ectopic expression of Sox2 was enough to turn mouse and human fibroblasts into induced neural stem cells (iNSCs) by culturing the transfected cells in neural differentiation medium supplemented with EGF and FGF2 (Ring *et al.* 2012). Aside from its role in pluripotency, Sox2 is also one of the main markers for neural stem cells, and has been shown to be a master regulator of neural fate. The fibroblasts reprogramed with Sox2 also expressed Nestin, another NSC marker, and gained morphology similar to iNSCs derived from iPSCs, gaining intrinsic expression of Sox2 without expressing any other pluripotency related genes and were capable of neuro-sphere formation. Neuro-spheres are floating agglomerates of neural cells and culturing these neuro-spheres is a common method for NSC expansion, although despite being a characteristic feature of NSCs they are not composed of NSCs alone, so it takes multiple rounds of neuro-sphere culture to isolate them (Ma *et al.* 2015). The obtained iNSCs showed capacity to differentiate into TUJ1⁺ immature neurons and later MAP2⁺ mature neurons in appropriate differentiation medium (Ring *et al.* 2012).

IV. Non-coding RNAs and their role in lineage commitment

There are various endogenous and exogenous molecules that intervene in signaling pathways involved in gene regulation. Among the endogenous intervenients, non-coding RNAs such as long non coding RNA (lncRNA) and micro RNA (miRNA) have been recently shown to possess preponderant functions in the regulation of pluripotency and cell fate related genes. LncRNAs are transcripts over 200 nucleotide long, transcribed from genes with no coding potential and so do not translate into proteins but many instead show regulatory functions themselves, and may act as enhancers of transcription in *cis* or repressors in *trans* by forming regulatory complexes with proteins, mRNA or DNA strands, or even by directly interacting with promoter regions (Johnsson *et al.* 2014). Many lncRNAs stand out due to their interaction with chromatin or with chromatin modifying protein complexes such as the Polycomb repressive complex (*PRC*), aiding methylation and acetylation processes that determine the silencing or activation of specific genes in the cell (Guttman *et al.* 2011).

A group of lncRNAs that are transcribed from intergenic regions known as large intergenic non-coding RNAs (lincRNA) have been reported to have an important contribution in the transcription of pluripotency related genes. LincRNAs are targets of various transcription factors and affect the cells global gene expression, especially in ESCs, where some of these were shown to influence the expression levels of key elements such as the TF Nanog, an important marker of pluripotency and a regulator of

pro-undifferentiation pathways (Guttman et al. 2011; Y. Wang et al. 2013). Therefore, lincRNAs are associated to the maintenance of pluripotency via suppression of differentiation pathways and their interaction with diverse protein complexes supposes they might possess cell-specific functions as flexible scaffolds in the formation of larger functional units (Guttman et al. 2011).

A crucial initial step in iPSC reprogramming is the mesenchymal-epithelial transition (MET), which is stimulated by the expression of Oct4 and Sox2. These TFs lead the transcription of a large number of regulatory RNAs, among them are the miRNAs of the miR200 family. Some RNAs of this family are required in the MET process, as in response to Oct4 and Sox2 they will act to suppress the Zeb1 and Zeb2 genes, as well as BMI1 and SUZ12 components of PRC1 and PRC2, respectively, allowing expression of the epithelial junction protein E-Cadherin to stabilize the endothelial state (Park et al. 2008; G. Wang et al. 2013). The proteins that rise from the Zeb (*Zinc finger E-box-binding homeobox*) family genes are also transcription factors and consist of multiple functional domains, two zinc-finger domains separated by a central homeodomain that are believed to repress gene expression by binding of the zinc-finger domains to the 5' regulatory regions of target genes, these sequences being coincident to DNA binding sites of basic helix-loop-helix TFs (Hegarty et al. 2015). Zeb2 expression is associated to a signaling cascade responsible for epithelial-mesenchymal transition (EMT), the reverse process of MET, where it acts as a repressor of E-Cadherin transcription (Vandewalle et al. 2005). Therefore, repression of Zeb2 during early iPSC reprogramming aids surpassing the mesenchymal-epithelial barrier, as performed by the mir200 family whose miRNAs bind directly to the 3' UTR of Zeb2 (G. Wang et al. 2013). Otherwise, Zeb2 upregulation stimulates EMT which is a common process of embryo development where Snail1 induces the expression of Zeb1/2. Snail1 influences Zeb2 expression by interfering with the processing of a long intron in the 5' UTR of the Zeb2 mRNA where there is an IRES (internal ribosome entry site) sequence that is essential for its translation into functional protein. Zeb2 expression is also regulated by the expression of a lincRNA transcribed from the antisense strand of Zeb2 named Zeb2-NAT, a natural antisense transcript (NAT) product that regulates Zeb2 expression in conditions of cell stress such as EMT by overlapping the splicing site of the intron where the IRES is present. This leads to the conservation of said IRES and therefore to translation of the processed mRNA to functional protein (Beltran et al. 2008). **Figure 1.3** shows how non-coding RNAs interplay with other main regulators during endothelial-mesenchymal and mesenchymal-endothelial transitions (Pauli et al. 2011).

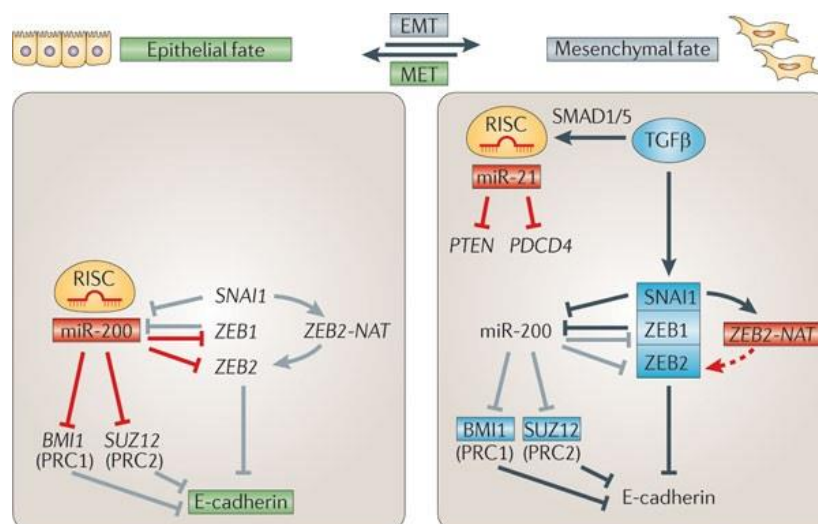


Figure 1.3 - Non-coding RNA regulation of mesenchymal-epithelial transition by controlling E-Cadherin expression. Zeb2NAT is induced by Snail1 promoting ZEB2 processing which will repress E-Cadherin leading to EMT (Pauli et al, 2011).

V. LncRNA regulation as a potential direct reprogramming enhancer

The possibility to obtain multipotent cells with fewer manipulation is promising enough, however reprogramming procedures face other limitations such as low reprogramming efficiency, particularly in iPSC reprogramming where efficiency usually doesn't get over 1% and though direct reprogramming has far greater efficiency, most methods still need improvement or optimization. Reprogramming efficiency is affected by both the cell line of origin and target cell line patterns, being that some cell lines prove to be more or less permissive to reprogramming than others, in regard to the target cell line of interest affecting the efficiency it is necessary to take in consideration that these variations may be related to the need for optimized reprogramming protocols that do not exist yet for most cell lines. Other limitation in terms of efficiency come from the age of reprogramming cells. Reprogramming cells from advanced age donors has shown to have lower efficiency than reprogramming fetal cells or cells from infant donors, the same way the reprogramming efficiency gets lower when reprogramming cultured cells with multiple passages. This correlation between age and reprogramming efficiency is most likely due to cellular senescence related phenomena (Trokovic et al. 2015). A possible way to enhance reprogramming of aged cells might come from lowering the cells primary barriers to the differentiation program applied, like MET in early iPSC reprogramming. Finding ways to facilitate this crucial steps may improve reprogramming efficiency, a process in which LncRNA may play an important role. For instance Zeb2-NAT regulates Zeb2 who plays a critical role in the reprogramming initiation, in this particular case for iPSC reprogramming (G. Wang et al. 2013). Due to the importance of transitions such as MET or EMT in the process of changing cell character and the recognized role of Zeb2 in said transitions (Goossens et al. 2011; Hegarty et al. 2015), Zeb2-NAT may prove to be a reliable tool in regulating his sense counterpart having indirect impact in the mentioned process. Transient down-regulation of this NAT during early reprogramming could pull down ZEB2 levels towards MET, what could diminish the cell's resistance to lineage transition facilitating pluripotency-related TF based reprogramming.

An easy and effective way to achieve this downregulation of lncRNAs is by targeting their transcripts with locked nucleic acid (LNA) based oligos. LNA oligos are oligonucleotides with at least one bicyclic furanose unit locked, this change confers them higher affinity and specificity for complementary sequences, making them perfect to target single stranded RNAs. LNA GapmeR™ are particularly interesting for downregulation, since these gapmers have modified LNA fragments flanking a 7-10nt DNA gap sequence and the heteroduplex formed by the hybridization between the antisense LNA oligo and target mRNA promotes RNase H mediated degradation, and consequently gene silencing (Kauppinen et al. 2005). LNA GapmeR™ action mechanism is represented in **Figure 1.4**.

Downregulating Zeb2NAT using LNA GapmeRs should demonstrate how interfering with this long non-coding RNA has influence on Zeb2 translation to functional protein, and therefore in the establishment of epithelial or mesenchymal character during reprogramming. By interfering with Zeb2 translation we interfere with the signaling cascade involved in mesenchymal fate definition, thus facilitating the early process of MET involved in iPS cell character shift, possibly increasing the population of epigenetically activated cells that will be reprogrammed, increasing direct conversion efficiency. Like Zeb2NAT, other antisense LncRNAs are recently being found to regulate their sense homologues. Getting a better understanding of how antisense noncoding genes interact with their sense counterparts and finding regulatory patterns could prove to be very important for targeting based therapies, since every day more lncRNAs are found to present function and many of them present associated pathologies.

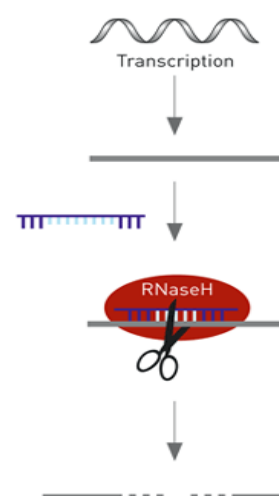


Figure 1.4 - Illustration of LNA GapmeR™ action (www.exiqon.com)

A massive amount of lncRNAs are still being discovered and characterized by high-throughput techniques such as RNA-seq and many of these newly characterized lncRNAs attract attention in regard to their possible or likely role in cell differentiation and cell fate commitment. In a recent study *Ramos et al* introduced an lncRNA called Pnky. This Pnky is a lincRNA that is polyadenylated and evolutionarily conserved, who is highly expressed in neural stem cells of the ventricular sub-ventricular zone of both human and mice adult brains. Downregulation of this 1560nt or 825nt long transcript in human or mice, respectively, appears to stimulate NSC differentiation with an increase in transit amplifying cell proliferation and neuronal fate commitment (Ramos et al. 2015). Pnky was found to interact with PTBP1, forming a complex who regulates neuronal repression by alternative splicing as illustrated in **Figure 1.5 A**, however the exact mechanisms of this regulation are still elusive. PTBP1 is an RNA-binding ribonucleoprotein of the PTB (*polypyrimidine tract-binding*) family, known to promote neuronal gene repression mediated by miRNAs. In fact knockdown of PTBP1 alone is sufficient to cause trans-differentiation of fibroblasts to neurons due to its effect on RE1-Silencing Transcription factor (REST), a restrictive complex that targets TFs involved in neuronal commitment. This RNA-binding protein is a target of miR124, a microRNA that targets components of the REST complex, in turn REST represses a large set of neuronal genes including miR124 itself, in a regulatory loop where PTBP1 acts as a negative regulator of miR124 (Xue et al., 2013; Kawahara, Imai, & Okano, 2012). **Figure 1.5 B** shows these regulatory interactions during NSC differentiation. PTBP1 is also a negative regulator of another PTB protein, PTBP2 who inversely to PTBP1 promotes the activation of neuronal genes and is necessary for the generation of neuronal precursors from NSCs (Ramos et al. 2015). However, Pnky and PTBP1 seem to act in an independent manner as knockdown of either one in NSCs has similar consequences in neuronal commitment without affecting the other's expression levels, and since Pnky is highly expressed in neural stem cells but not in other cell lines like fibroblasts, where PTBP1 also plays its repressive role, it would be of interest to explore if this lincRNA has any further contribution to multipotent neural cell regulation.

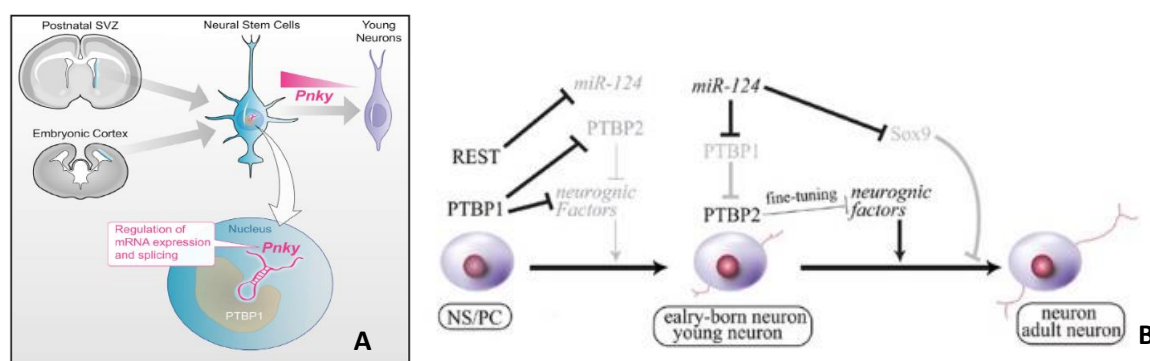


Figure 1.5 - A Representation of Pnky localization and influence in neurogenesis from NSCs (*Ramos et al 2015*); **B** Schematic representation of neuronal fate commitment regulation by PTB proteins. REST and PTBP1 repress neuronal differentiation of neural stem/progenitor cells, while PTBP2 promotes neuron maturation (*adapted from Kawahara et al 2012*)

A point that favors such exclusive role is that Pnky expression seems to be lost in GFAP (*glial fibrillary acidic protein*) negative cells, and it's not expressed in neural stem cells of the dentate gyrus. (Ramos et al. 2015) This might be an indicator that along with its role in repressing neurogenesis with PTBP1, Pnky long noncoding RNA may have some contribution to other NSC properties, possibly the maintenance of some glial character by complementary positive regulation of glial-related genes that along with the repression of neuronal genes give NSCs their overall expression pattern, regardless or not of PTBP1 presence.

The increased knowledge about lncRNA and how they contribute to gene regulation could provide other reprogramming approaches, for example, by forcing the expression of lineage specific lncRNA in a similar fashion to what *Yamanaka et al* did with transcription factors, or co-expression of TFs along with lncRNAs could possibly improve reprogramming efficiencies. In iNSC reprogramming,

Pnky is a consistent candidate for such approach, since its loss pushes NSCs towards neuronal lineage commitment, so stimulating cells towards a neural program while simultaneously refraining definitive neuronal commitment could lead to establish an intermediate progenitor-like state.

Not only Pnky, but inclusively Zeb2-NAT could make a good candidate for NSC reprogramming later on by this approach, since it's a positive regulator of its sense counterpart Zeb2, who has also been shown to play an important role in early nervous system development (Hegarty et al. 2015). As a Smad interacting protein, ZEB2 is involved in a lot of developmental processes, since Smad proteins are effectors of the TGF β /BMP pathway, and these proteins are essential for embryonic development, inclusively the early nervous system. Zeb2 is required for neuroectoderm formation after embryogenesis, where Smad-inhibition, Sox2 and Fgf induce Zeb2 expression, who is also preponderant in brain development, where its loss leads to pathogenic traits such as Mowat-Wilson syndrome, a genetic disease with variable penetrance caused by a mutation on one of the Zeb2 alleles that ultimately leads to severe mental retardation, microcephaly and seizures in affected patients. ZEB2 neural induction is dependent on the interaction with CtBP (*C-terminal Binding Protein*) and NuRD (*Nucleosome Remodeling Deacetylase*) co-repressor complexes, acting together to repress BMP/Smad signaling. Its role in EMT is also very important for neural crest cell (NCC) migration and differentiation during neural tube formation (Hegarty et al. 2015). **Figure 1.6** shows a representation of Zeb2 multi-functions in neuroectoderm formation and NCC maturation. Although NCCs and NSCs are very distinct, the fact that ZEB2 is able to induce neural genes makes Zeb2-NAT a possible intermediate in the neural program.

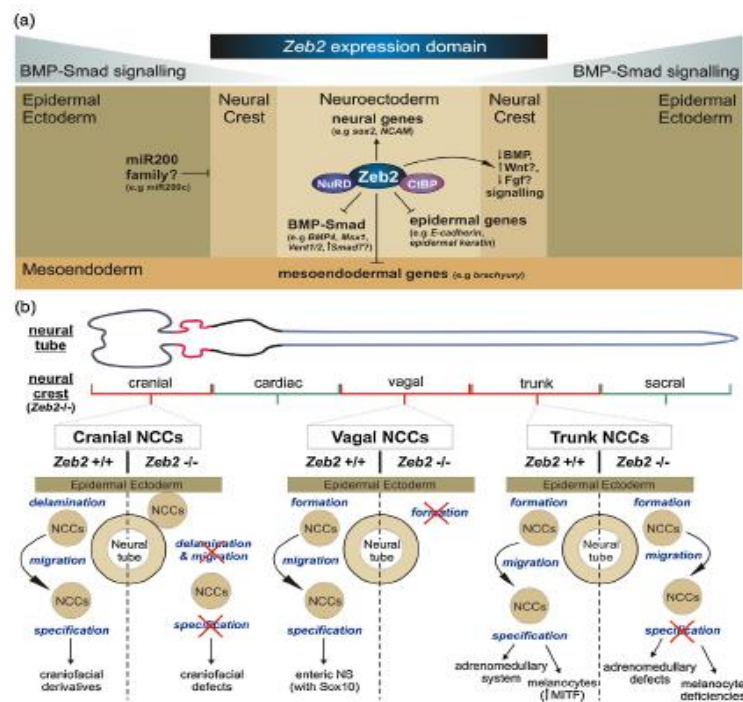


Figure 1.6 – (a) Representation of Zeb2 role in neuroectoderm formation by inducing neural genes and repressing epidermal and mesoendodermal fate (b) Schematic representation of Zeb2 knockout in cranial, vagal and trunk neural crest cell populations effect on NCCs formation, migration and specification (Hegarty et al, 2015)

As mentioned before, another important barrier to cellular reprogramming is aging. The limiting alterations to reprogramming that occur during the aging process may be caused by stochastic events such as mutations in the genome due to exposure to environmental or working conditions of the individual during his lifetime, or part of a defined aging program (Ashapkin et al. 2015). At the epigenetic level, alterations occurring with aging lead to the activation or suppression of specific pathways, some of which might become barriers of reprogramming. On the other hand, the

reprogramming procedure itself has been shown to affect these epigenetic alterations, specifically iPSC reprogramming to a *naïve* state can erase some of the epigenetic marks gained throughout life, resetting the cell to the embryonic stage even at the epigenetic level (Ashapkin et al. 2015). This epigenetic reboot of the cell is, however, limited to alterations that are programmed for the organismal aging as a whole and does not affect alterations that resulted from stochastic events, such as point mutations. Alternatively, the trans-differentiation approach does not need so profound modifications and seems not to involve epigenetic recovery of any kind, so cells reprogramed directly from one lineage to another will retain the aging associated epigenetic alterations they got (Yang et al. 2015). Nonetheless, how direct conversion using pluripotency related TFs will affect aging marks is still poorly understood.

In this work we sought to find ways to use lncRNAs to condition, improve or substitute pluripotency TF mediated cell reprogramming, particularly in aged cells. Our first approach to increase reprogramming efficiency was based on the need for the initial MET step that occurs in iPSC early reprogramming, where targeting the lncRNA Zeb2NAT with LNA GapmeRs proved to facilitate this step by downregulating the Zeb2 gene and possibly increase the epigenetically activated cell population that will be reprogrammed by direct conversion, leading to higher yields. Adapting and optimizing the protocols to combine retroviral expression of TFs and LNA targeting was our first priority, but also finding lncRNAs that act on regulatory pathways possibly needed for direct conversion to occur, hence an alternative lncRNA overexpression approach is also under consideration. Considering all this aspects, the progress of this work so far is presented as follows.

2. Materials and Methods

Reprogramming protocols

Human fibroblast (HuF) cell lines of various ages were transduced with a single retrovirus containing the hOct4 or hSox2 genes for the iHPC and iNSC protocol, respectively, and GFP expressing retrovirus was used as a retroviral transfection control. LNA Gapmers (Exiqon) were used to target Zeb2-NAT transcripts for downregulation. All Kit based procedures were performed according to the suppliers protocols. All retroviral vectors used in the reprogramming experiments presented were a kind gift from Manuel Serrano, CNIO, Madrid (plasmid maps and designs are available at Addgene website (www.addgene.org)).

Retroviral production

4,5x10⁶ HEK 293T cells were plated in gelatin coated 10 cm plates and transfected with 4µg pMXs-hOCT3/4 (Addgene #17217) or pMXs-hSOX2 (Addgene #17218) plasmids using X-tremeGENE™ 9 DNA Transfection Reagent (Roche) and 4µg AMPHO (NBP2-29541) packaging plasmid. 293T cells were left in DMEM+10%FBS overnight. The next day culture medium was replaced and cells were left for retrovirus production. 36 hours after the transfection, retroviral containing medium were collected every 12h for a 48h period. Collected retroviral medium was mixed with polybrene 1:000 v/v and used directly to infect fibroblasts in 24 wells plate using 500µL of viral medium per well. As a positive control, the same protocol was used with the retroviral plasmid pBabe-GFP and applied to each HuF cell line used to access transduction efficiency.

Zeb2-NAT downregulation

Downregulation of Zeb2-NAT in human fibroblasts was performed using 2 LNA-GapmeRs (EXIQON™) targeting specifically Zeb2 antisense transcripts in two transfections, the first 24 hours before the first retroviral infection and the second 24 hours after the last retroviral infection. Transfection with a scramble LNA with no target was performed as negative control. LNA transfections were performed using the Lipofectamine® RNAiMAX Reagent (Thermo Fisher Scientific) according to the supplier's protocol. Each transfection was performed in triplicate for every HuF cell line in each experiment.

Human cell lines of different ages used in reprogramming experiments are presented in **Table 2.1**.

Table 2.1 - Brief description of human cell lines used for reprogramming experiments.

<i>Cell line</i>	<i>Description</i>	<i>Origin</i>		<i>Suppliers reference</i>
WI-38	Human fibroblasts	fetal	Lung	ATCC No CCL-75
3YR	Human fibroblasts	skin	3 year old donor	GM05565
11YR	Human fibroblasts	skin	11 year old donor	GM00323
73YR	Human fibroblasts	skin	73 year old donor	AG06952

Culture Conditions

a) Fibroblast culture

All HuFs were maintained on 10cm dishes in DMEM (GIBCO) w/ 15% FBS, 2mM L-Glutamine and Pen-Strep (penicillin at 100 units.mL⁻¹ and streptomycin at 100µg.mL⁻¹) (P4333 Sigma) prior to the reprogramming experiments.

b) iHPC

Reprogramming of human fibroblasts was performed based on the protocol of *Szabo et al*, except retrovirus containing the hOct4 transcription factor gene for overexpression were used instead of lentivirus. 2.5x10⁵ HuFs per well were plated on Matrigel™ (BD Biosciences 354277) coated 24 well plates. 2 separate experiments were performed using 2 HuF cell lines of different age divided in 12 wells per plate. Fibroblasts were cultured in DMEM + 15%FBS until retroviral infection. After transduction with hOct4 and LNA treatment, medium was changed to DMEM F12 medium (GIBCO 31330038) supplemented with 10% Knock-out serum replacement (GIBCO 10828028), 1% non-essential amino acids, 1mM L-glutamine, 0.1mM β-mercaptoethanol, 16 ng.mL⁻¹ FGF2 (R&D Systems 234-FSE-025) and 30 ng.mL⁻¹ IGF2 (Millipore 01-142), for 21 days.

c) iNSC

Reprogramming of human fibroblasts was conducted according to the procedure by *Ring et al* by retroviral infection with retrovirus expressing the human pluripotency transcription factor Sox2. Retroviral infection and LNA treatment were performed the same as for the iHPC procedure. Transduced HuFs were cultured in 24 wells plates coated with glass coverslips over gelatin and mitomycin C treated Feeder cells, in ReNcell medium (Millipore) supplemented with 20 ng.mL⁻¹ of human FGF2 (R&D Systems) and EGF (Millipore GF144). Medium was replaced every day for 10 days.

Reprogrammed cells analysis

a) iHPC

After 21 days in culture, fibroblasts from each well were individually detached with trypsin, washed by centrifugation, suspended in PBS with 2% FBS and incubated for 20min in the dark at 4° C with 5µL of APC mouse anti-human CD45 antibody (BD Pharmingen™ 560973) per sample for staining. Stained cells were washed again to remove residual antibody and suspended in PBS with 2% FBS once again. Fibroblasts expressing the CD45 hematopoietic marker were counted by flow cytometry using in the BD Accuri C6 flow cytometer (BD Biosciences™).

b) iNSC

Reprogrammed HuFs were directly collected for RT-qPCR analysis to check for expression of neuronal and neural stem cell markers without neuro-sphere formation. After 10 days in differentiation medium, cells were collected for total RNA extraction using PureZOL™ RNA Isolation Reagent (Bio-Rad 7326880). Every RNA sample corresponding to a well of the plate was treated with DNase for purification and 500ng of RNA were used for cDNA synthesis of each sample with the Transcriptor

High Fidelity cDNA Synthesis Kit (Roche 05091284001). The obtained cDNA was then analyzed by RT-qPCR with iTaq™ Universal SYBR Green Supermix (Bio-Rad 172-5124) for Nestin, Sox2, Oct4, Tuj1, Nurr1, MAP2, PAX6, GFAP and Zeb2NAT levels (for primer sequences see supplementary data) using ViiA 7 System or 7500fast RT-qPCR equipment from Applied Biosystems™.

Pnky cloning

mPnky

The brain ventricular zone of a 12 week-old C57BL/6 mouse was dissected and homogenized. Total RNA was then isolated using PureZOL™ RNA Isolation Reagent, precipitated with isopropanol and DNase treated with recombinant DNase I, RNase-free (Roche) and cDNA was produced by reverse transcription of the total RNA extracted, using the Transcriptor High Fidelity cDNA Synthesis Kit (Roche). Pnky lncRNA was cloned by PCR with the Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific F530L) using Forward and Reverse primers with 100% homology to the 3' and 5' ends of Pnky (for primer sequences used for PCR reactions see supplementary data). PCR reactions were performed using MyCycler Thermocycler (Bio-Rad #1709713) and after optimization of cycle time and annealing temperature, optimized PCR conditions used for mPnky amplification were as follows:

1st Cycle - Melting: 1' at 98° C

2nd Cycle - Melting: 20'' at 98° C Annealing: 30'' at 58° C Extension: 60'' at 72° C (x35)

3rd Cycle - Extension: 10' at 72° C

The resulting PCR products were separated by electrophoresis on 1% TAE agarose gel stained with Midori Green Advance (Nippon Genetics Europe). Obtained bands with 800-900 bp were extracted from the gel and purified by NZYGelpure Kit (NZYTech MB01101). DNA samples were quantified using NanoDrop 2000 (Thermo Scientific) and sent for Sanger sequencing. From the sequencing results we identified mPnky among the samples by BLAST analysis (blast.ncbi.nlm.nih.gov/Blast.cgi).

Pnky retroviral plasmid construction

a) pMXs Backbone (sticky ends)

Mouse Pnky was amplified by PCR with the Phusion High-Fidelity DNA Polymerase using primers with a homology sequence and a restriction site for BstXI (NEB R0113S) to make sticky ends for directed ligation to the vector. Vector plasmid was obtained by excision of mOct3/4 from pMXsOct3/4 (Addgene #13366) plasmid with BstXI restriction enzyme. Vector plasmid was dephosphorylated and ligated to the restricted Pnky insert using the Rapid Dephos & Ligation Kit (Roche 04898125001). Ligations were performed using 1:1, 1:3 and 1:5 vector/insert ratios, using the dephosphorylated plasmid alone as a negative control. After ligation, DH5α competent bacteria were transformed with the ligated and control plasmids mix and plated in Lysogeny broth (LB) + agar plates with ampicillin 0,1% for clone selection and incubated at 37° C overnight. Grown colonies were picked and grown in 5mL LB medium with ampicillin 1:1000 v/v in a rotation incubator at 37° C and 220rpm overnight. The plasmids were extracted and purified from the bacterial pellet using GeneJET Plasmid Mini-prep Kit (Thermo Fisher Scientific K0503). Extracted plasmids were then digested with FastDigest EcoRI (Thermo Fisher Scientific FD0274) restriction enzyme, who cuts mPnky at 167bp but not the

plasmid. Linearized plasmids were sent for Sanger sequencing to confirm Pnky insertion. All reagents and kits were used following the suppliers protocols.

b) pBabe Backbone (blunt ends ligation)

Despite the effort, mPnky could not be inserted into the pMXs plasmid, so we changed the approach to blunt end ligation and later to a different retroviral plasmid backbone. This time, pBabe-GFP plasmid was cut with HindIII (Thermo Fisher Scientific ER0501) and NheI (Thermo Fisher Scientific ER0971) restriction enzymes, according to the supplier's protocol, using 1X Tango Buffer and 2-fold excess of HindIII and incubating at 37° C overnight. Enzymatic restriction excised the GFP ORF. After restriction, plasmid was treated with Klenow Fragment (Thermo Fisher Scientific EP0054) enzyme for plasmid vector blunting by fill-in of the 5'-overhang and 3'-overhang loss, according to the supplier's protocol. Total reaction volume was run on TAE 0.8 % agarose gel to separate bands by electrophoresis and the higher band corresponding to the pBabe empty vector was excised from the gel and purified using the NZYGelPure Kit. Pnky insert was treated with T4 Polynucleotide Kinase (Thermo Fisher Scientific EK0031) to guarantee that Pnky DNA obtained from PCR reaction was phosphorylated before ligation reaction was performed according to the supplier's protocol. Once purified, the plasmid was dephosphorylated and ligated to the Pnky insert with Rapid Dephos & Ligation Kit (Roche). Transformation on DH5 α competent bacteria and plasmid amplification and extraction were performed as described before for the pMXs plasmid.

3. Results and discussion

Zeb2NAT in early reprogramming

Earlier studies by *Bernardes de Jesus et al* at IMM showed that transient downregulation of Zeb2NAT during early reprogramming of adult fibroblasts to iPS cells improves the reprogramming efficiency by facilitating the initial MET process. To access if the role of Zeb2NAT in regulating Zeb2 during the early stage where MET has to occur for fibroblasts to reprogram into iPSCs is common to direct reprogramming of human fibroblasts with a single pluripotency associated transcription factor, a similar approach was performed in two different reprogramming protocols. One reprogramming protocol for iHPCs based on *Szabo et al* reprogramming procedure, and another for iNSCs based on *Ring et al* reprogramming protocol. Transient downregulation of Zeb2NAT was performed using LNA GapmeRs to target human Zeb2NAT transcripts. HuFs used in each experiment were transfected with LNAs after retroviral transduction on both iHPC and iNSC protocols. To access Zeb2NAT downregulation efficiency of the GapmeRs, RT-qPCR was performed from total RNA extracted from HuFs transfected with LNAs after 15 days in culture, results are shown in **Figure 3.1**. Zeb2NAT expression levels decrease 50-80% compared to wild type HuFs.

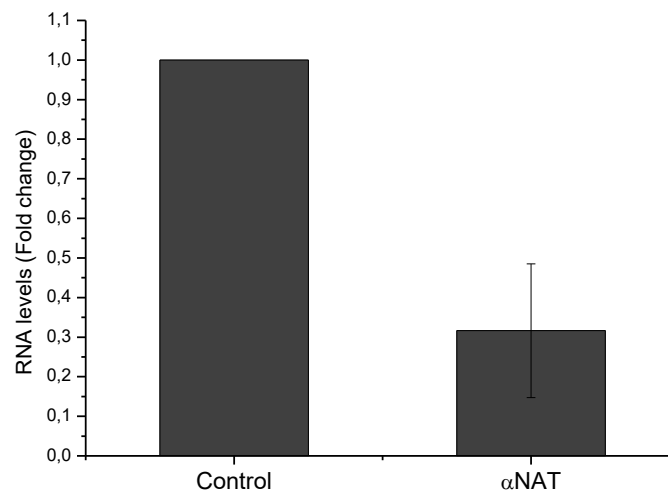


Figure 3.1 - Downregulation efficiency assay, Zeb2NAT expression level in WI-38 HuFs treated with LNA GapmeRs targeting Zeb2NAT transcripts (αNAT) compared to wild type WI-38 HuFs (Control).

To access infection efficiency of retroviral plasmids used for cell reprogramming, retrovirus containing the GFP expression gene were used as positive control. Each HuF cell line used in the experiments was infected using viral DMEM medium containing viral particles with pBabe-puroGFP expression plasmid and transduction efficiency was accessed by microscopy. Infected cells expressed the green fluorescent protein as seen in **Figure 3.2**. Different age cell lines showed different infection efficiency, with older cell lines showing lower GFP expression. This is most likely due to the fact that retrovirus only infect multiplying cells and the older the HuFs, the lower their division capacity.

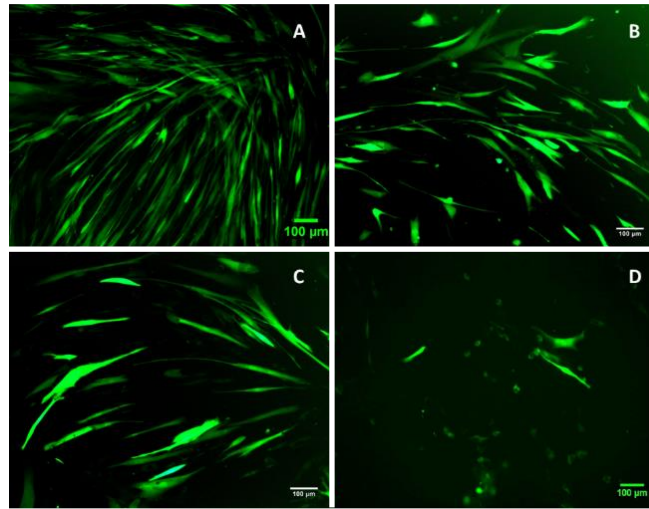


Figure 3.2 - Human fibroblast cell lines WI-38(A), 3YR (B), 11YR (C) and 73YR (D) transduced with GFP retrovirus for transduction efficiency positive control (Photos were taken with Zeiss Axiovert 200M microscope with x10 amplification)

After testing transfection efficiencies, we started reprogramming assays. First experiments were performed using all of the mentioned HuF cell lines, however due to the identified transduction efficiencies identified by the GFP expression assay, some lines were excluded from the experiments as reprogramming was not effective due to low retroviral infection and consequent low expression levels of the transcription factor of interest. Cell lines used for each experiment were chosen considering protocol limitations and adequate age HuF cell lines more suited for the experiments were determined.

Since Zeb2NAT is only involved in the early steps of reprogramming, the protocols tested were only applied until a checkpoint defined at the rise of the first differentiation markers, after which cells were collected and tested for those specific marker expression and no expansion methods were applied at this point.

iHPC

Reprogramming of human fibroblasts with Oct4 was performed in two 24 well plates, each plate split in 12 wells of each HuF cell line. Cells were cultured under 5 different conditions, which were transduced with Oct4 alone, with Oct4 and Zeb2NAT LNA GapmeRs, Oct4 and Scramble LNA GapmeRs with no target, in differentiation medium, and Oct4 in common fibroblast medium (DMEM+15% FBS and L-glutamine) and a well with HuFs transduced with GFP alone for control. The experiment model is presented in **Figure 3.3**.

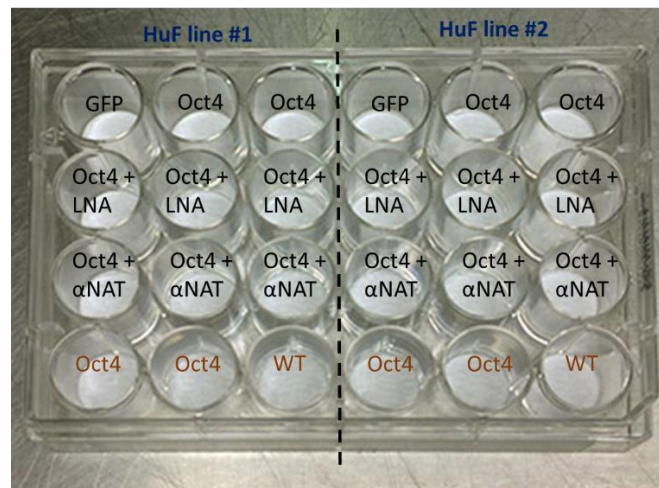


Figure 3.3 - Experiment model for fibroblast reprogramming with Oct4 (Oct4 - HuFs transduced with hOct4; LNA - HuFs treated with scramble LNA; α NAT - HuFs treated with LNA GapmeRs targeting Zeb2NAT; GFP - HuFs transduced with GFP; WT - Wild type HuFs). Conditions in black are for HuFs cultured in differentiation medium and in brown for fibroblast culture conditions during 21 days after retroviral transduction and LNA treatment.

The four HuF cell lines were split in 2 plates, the first with WI-38 and 73YR and the second with 3YR and 11YR cell lines. However due to poor retroviral efficiency, as seen by the GFP assay, only the WI-38 and 3YR cell lines were able to grow and give rise to colonies under differentiation conditions. Although the reprogramming protocol was tested in all cell lines, it is important to mention that these two plates are part of two separate experiments and are referred to by chronology of said experiments, and some methodological corrections were made from one to the other.

In the first experiment, which relates to the WI-38 and 73YR cell lines, the differentiation medium applied was only supplemented with IGFII and FGF2, as the original reprogramming protocol by *Szabo et al* shows that to be enough for CD45⁺ colony formation although with low yields. To access the capacity of HuFs to reprogram under this conditions the youngest and the oldest cell lines were used in this first approach. 73YR HuFs completely failed to reprogram, as cell multiplication is slow in this cell line and retroviral infection needs dividing cells to occur. WI-38 HuFs on the other hand showed clear morphology change under differentiation conditions, as observed on the photos in **Figure 3.4** these changes started with the appearance of colonies of smaller cells, and although most wells treated with Oct4 retroviral medium had grown colonies, size and number of colonies seems to differ under LNA treatment conditions. Considering that aside from the supplementation, the culture conditions applied are the same used for iPSC reprogramming, the morphology change observed doesn't necessarily mean that cells are activating the hematopoietic program. The alterations observed could most likely be just the consequence of partial reprogramming mediated by the expression of Oct4.

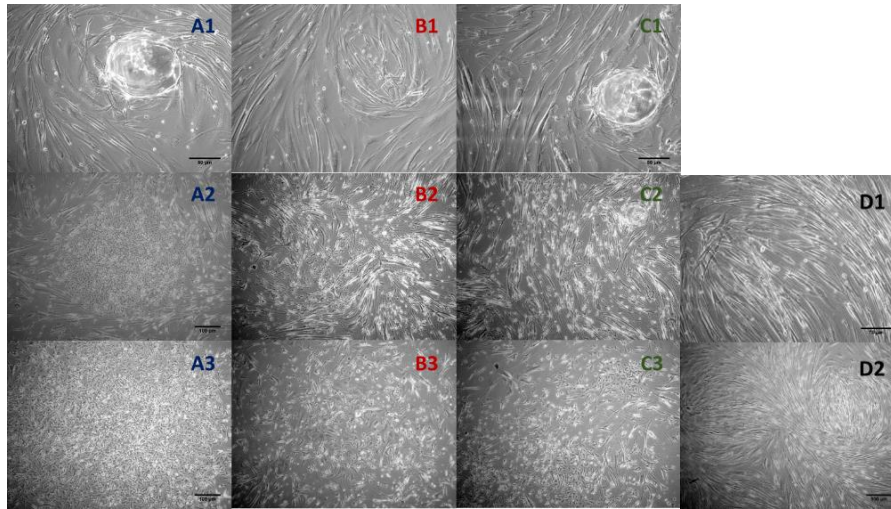


Figure 3.4 - WI-38 human fetal fibroblasts transduced with Oct4 alone (A), with Oct4 and treated with Zeb2NAT LNA GapmeRs (B) or scramble LNAs (C) at day 5 (1), day 10 (2) and day 19 (3) in differentiation medium composed of DMEM F12 + 10% KSR medium supplemented with IGF2 and FGF2. WI-38 cells transduced with GFP alone as control at day 5 (D1) and day 19 (D2) also in differentiation medium (photos taken in Zeiss PrimoVert Microscope, amplification x4 (1) and x10 (2, 3)).

To determine the amount of cells that developed the pan-leukocyte marker CD45, transduced fibroblasts were collected after 21 days in differentiation medium, washed and stained with anti-human CD45-APC antibody and analyzed by FACS. Cells from mouse spleen were used as positive control for CD45 expression. The shift characteristic of CD45⁺ cell's profile obtained from the positive control when stained with antibody is presented in **Figure 3.5**, as well as the unstained and WT human fibroblast profile in FACS analysis for negative control.

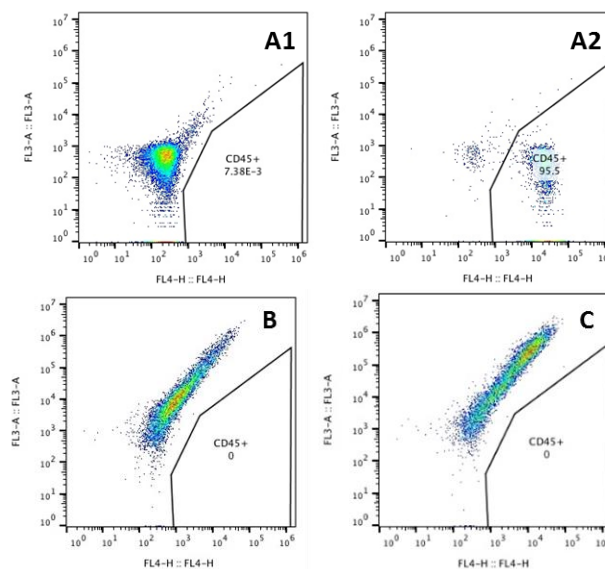


Figure 3.5 – FACS profile controls. CD45 positive control cells unstained (A1) and stained with anti-humanCD45 antibody (A2), HuFs transduced with Oct4 unstained (B) and WT HuFs stained for CD45 (C) for negative control.

Despite the clear appearance of colonies with different morphology from the original WI-38 fibroblasts, FACS analysis didn't show a considerable shift in stained cell's profile, as observed in **figure 3.6**. This result indicates that supplementing the medium with IGF2 and FGF2 alone when using retrovirus might not be stimulation enough to promote the rise of CD45⁺ cells, or at least CD45 expression is not high enough to be clearly identified by FACS.

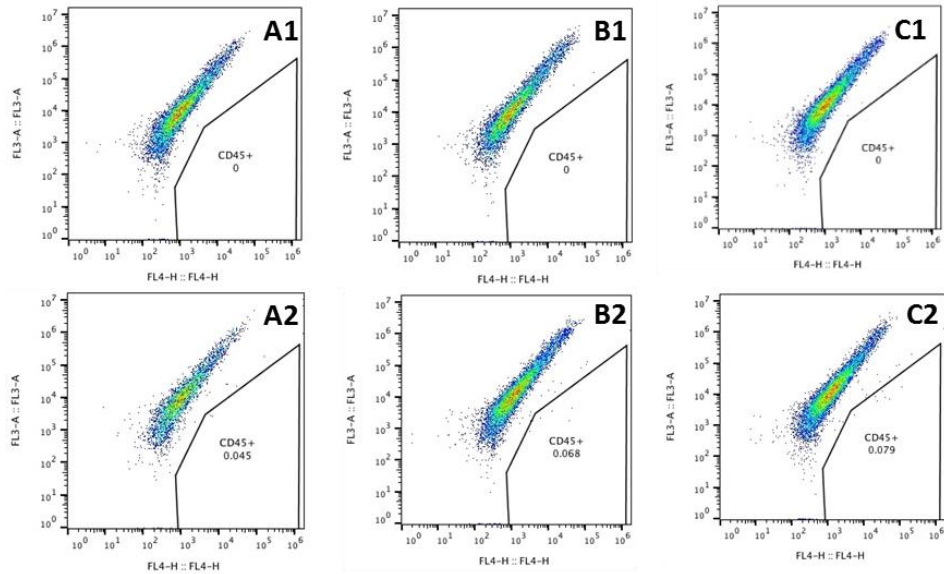


Figure 3.6 – WI-38 HuFs FACS profile when cultured in KSR medium supplemented with IGF2 and FGF2 transduced with Oct4 alone (**A,B** and **C**) and treated with Zeb2NAT LNA GapmeRs (**B**) or scramble LNAs (**C**) and stained for CD45 (1-Unstained, 2-Stained for CD45). FACS analysis performed using BD Accuri C6 with a cell count of 10000.

A second experiment was performed using the 3YR HuF cell line using the same experiment design, but this time the differentiation medium was supplemented with SCF and FLT3-LG at $300\text{ng}\cdot\mu\text{L}^{-1}$ in addition to the $30\text{ng}\cdot\mu\text{L}^{-1}$ IGF2 and $16\text{ng}\cdot\mu\text{L}^{-1}$ of FGF2 already used in the first experiment after day 10 in differentiation culture. This extra supplementation resulted in the rise of bigger colonies in all conditions, however cells transfected with Zeb2NAT LNA GapmeRs still showed smaller colonies when compared to Oct4 alone and Oct4 with Scramble LNAs, as observed in **Figure 3.7** (an amplified photo can be seen in supplementary data).

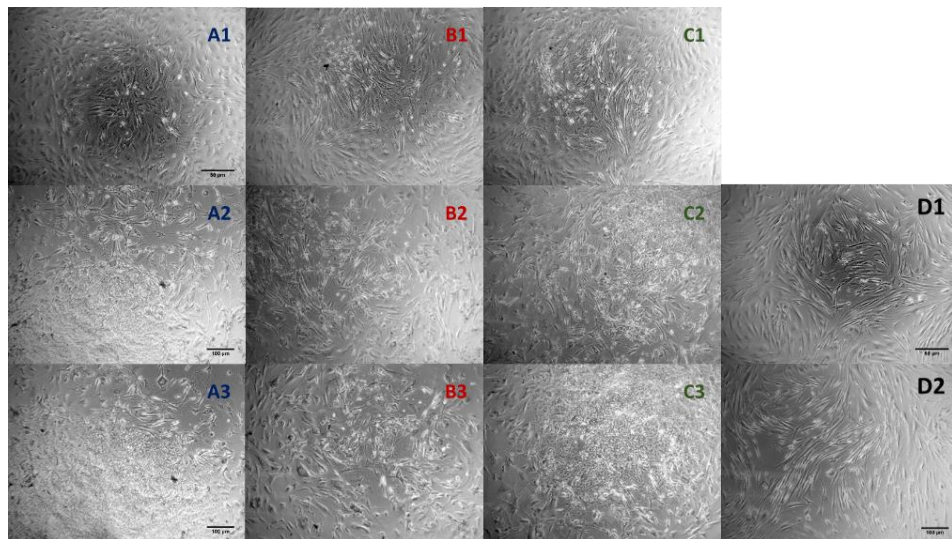


Figure 3.7 – 3YR human fibroblasts transduced with Oct4 alone (**A**), with Oct4 and treated with Zeb2NAT LNA GapmeRs (**B**) or scramble LNAs (**C**) at day 6 (**1**), day 14 (**2**) and day 21 (**3**) in differentiation medium composed of DMEM F12 + 10% KSR medium supplemented with IGF2 and FGF2. 3YR cells transduced with GFP alone as control at day 6 (**D1**) and day 21 (**D2**) in the same medium (photos taken with Zeiss PrimoVert Microscope, amplification x4 (**1**) and x10 (**2, 3**)).

Again, cells were collected at day 21 for FACS analysis. This time, there was a clear shift in the cell distribution profile of HuFs transduced with Oct4 (for cell distribution profile see Supplementary data) when stained with anti-humanCD45 antibody, and near 10% of counted cells were CD45⁺ cells. When downregulating Zeb2NAT, CD45⁺ cell number seems to drop almost 2 fold, as can be observed from **Figure 3.8**.

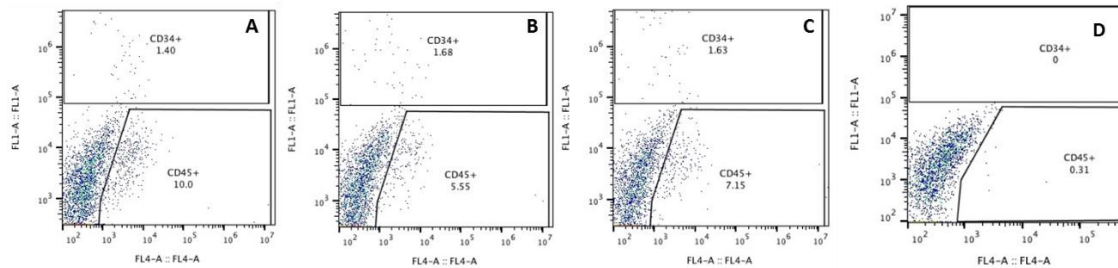


Figure 3.8 – 3YR HuFs cultured in KSR medium supplemented with IGF2 and FGF2 transduced with Oct4 alone (A,B and C) and treated with Zeb2NAT LNA GapmeRs (B) or scramble LNAs (C) and stained for CD45 and CD34. 3YR HuFs transduced with GFP retrovirus alone and cultured under the same conditions were also stained for negative control (D). Cell count=5000.

From the preliminary results presented, it appears that transient Zeb2NAT downregulation not only doesn't help the reprogramming process but might directly interfere with colony growth and CD45⁺ appearance from Oct4 transduced HuFs. However, more experimental replicates need to be performed to confirm this correlation.

Since the observed cell colonies showed clear distinct morphology from fibroblasts, staining with anti-human CD34 antibodies was also performed with the 3YR cell samples to check for the presence of the CD34 hematopoietic marker. As also shown in **Figure 3.8**, FACS analysis shows that a small percentage of cells seems to express the CD34 marker in Oct4 transduced cell samples, but there seems to be no co-expression of both hematopoietic markers. Still, the CD34⁺ population is very small and doesn't seem to be affected by Zeb2NAT downregulation. Unfortunately, at this point we did not have a positive control available, so further experiments need to be performed to confirm this trend.

iNSC

Human fibroblasts reprogramming to neural stem cells was performed according to *Ring et al* protocol consisting on the overexpression of the stem cell transcription factor Sox2 in neural differentiation medium (see materials and methods). Like the first experiment performed for the iHPC approach, the first experiment with this protocol was also performed using the HuF cell lines WI-38 and 73YR. Unlike in iHPC, 73 year old donor fibroblasts grew well under differentiation conditions, and showed morphology changes similar to the ones seen with WI-38. However, 73YR cells showed greater morphology heterogeneity in culture conditions, and after 9 days in ReNcell medium supplemented with EGF and FGF2, cells with clear neuronal morphology were observed (see supplementary data), particularly in the absence of Sox2 retroviral transduction but also, to a lesser extent, when Sox2 transduction is accompanied by Zeb2NAT down-regulation.

After the 10 day differentiation period cells were collected and total RNA extraction was performed for RT-qPCR analysis of neural markers under every condition showed in the experiment design presented in **Figure 3.9**. As before, GFP expressing fibroblasts were used as positive control for transduction efficiency and cultured under the same conditions as Sox2 transduced HuFs, working as negative control for iNSC reprogramming.

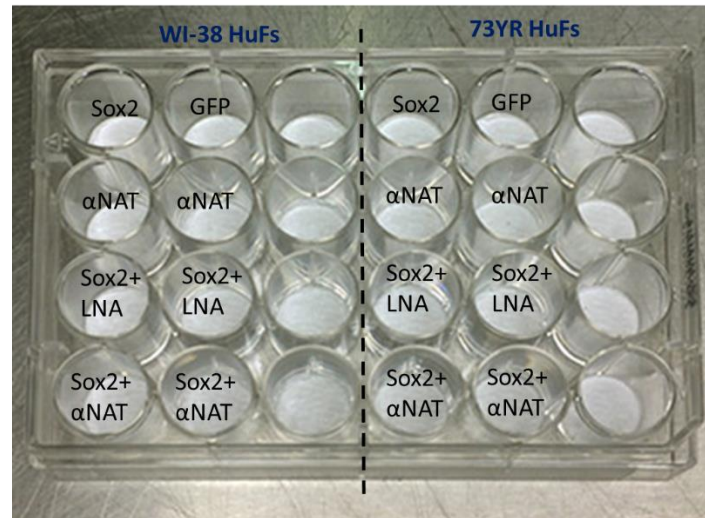


Figure 3.9 – Experimental model for fibroblast direct reprogramming with Sox2. All HuFs in this experiment were cultured for 10 days in ReNcell medium supplemented with EGF and FGF2 at 20ng.mL^{-1} , after retroviral transduction and LNA treatment (Sox2 - HuFs transduced with hSox2; LNA - HuFs treated with scramble LNA; α NAT - HuFs treated with Zeb2NAT targeting LNA; GFP - HuFs transduced with GFP retrovirus alone).

HuFs transduced with Sox2 showed morphology changes characteristic of iNSC reprogramming as described by *Ring et al*, defined by the formation of networks under differentiation conditions, these changes are presented in the photos of **Figure 3.10 A** and **3.11 A** for WI-38 and 73YR HuF cell lines, respectively.

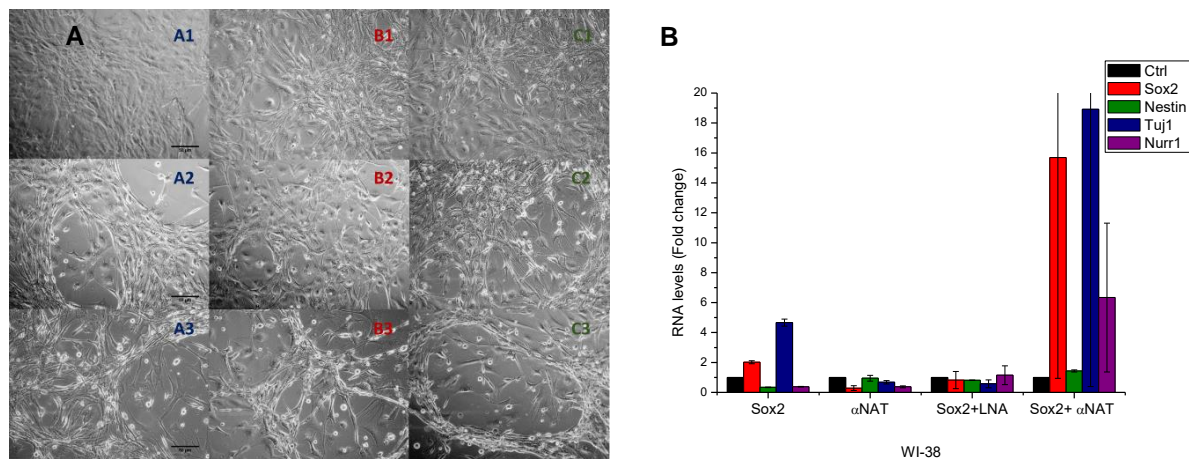


Figure 3.10 – **A** WI-38 HuFs transduced with Sox2 alone (A), with Sox2 and treated with LNA GapmeRs targeting Zeb2NAT (B) and with scramble LNAs (C) after 3 days (1), 5 days (2) and 9days (3) in culture with neural differentiation medium (for control cells see Figure D in supplementary data); **B** RT-qPCR result expression levels, presented as $\Delta\Delta C_t$ normalized using GFP retrovirus only transduced cells as control and GAPDH as housekeeping gene (for results using Actin as housekeeping see supplementary data) (photos taken in Zeiss PrimoVert Microscope, amplification x10).

RT-qPCR of all samples was performed to determine expression levels of Sox2, Nestin, Tuj1 and Nurr1. Sox2 and Nestin to access iNSC development and Tuj1 and Nurr1 to check for neuronal differentiation under culture conditions. RT-qPCR results obtained from WI-38 and 73YR reprogrammed cell samples are presented in **Figure 3.10 B** and **3.11 B**, respectively.

From this results it appears that LNA treatment is hindering Sox2 retroviral transduction, as Sox2 levels drop when LNA GapmeRs are used along with retroviral infection. This could mean that combining retroviral infection with LNA treatment still needs to be corrections to the protocol, possibly by increasing the time gap between infections to lower cell stress.

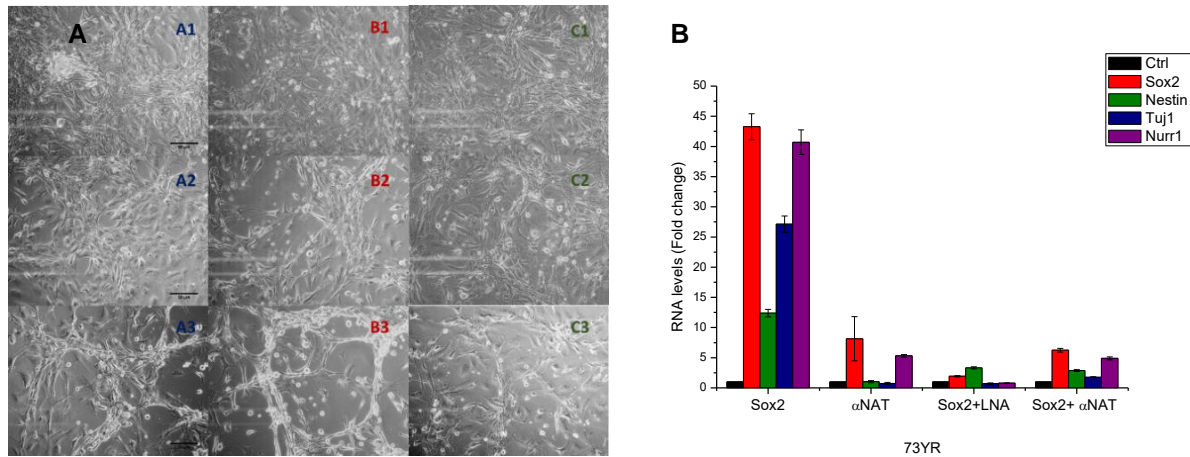


Figure 3.11 – A 73YR HuFs transduced with Sox2 alone (A), with Sox2 and treated with LNA GapmeRs targeting Zeb2NAT (B) and with scramble LNAs (C) after 3 days (1), 5 days (2) and 9days (3) in culture with neural differentiation medium (for control cells see Figure D in supplementary data); **B** RT-qPCR result expression levels, presented as $\Delta\Delta C_t$, normalized using GFP retrovirus only transduced cells as control and GAPDH as housekeeping gene (for results using Actin as housekeeping see supplementary data) (photos taken with Zeiss PrimoVert Microscope, amplification x10).

Some results don't match the morphology changes seen, this is probably due to a great variation in the levels of the housekeeping genes observed in the amplification C_t s, where values show variations up to 5 cycles for Actin and 8 for GAPDH, compromising a throughout evaluation of registered expression levels. This problem could be consequence of stress induced by HuF transduction with both retrovirus and LNA, therefore causing changes in the expression pattern of commonly expressed genes, or possibly due to RNA sample dilution after cDNA synthesis. Although some trends can still be observed from the data, unfortunately, combining the reprogramming protocol with LNA treatment still needs further optimization.

LncRNA overexpression by retroviral expression

Long noncoding RNA levels influence differentiation processes through diverse expression patterns. To access the possibility of improving reprogramming thru forced expression of cell line specific lncRNA, we proceeded with the cloning of the Pnky lncRNA, which is commonly expressed in NSCs, as seen in **Figure 3.12** and plays a role in regulating NSC differentiation to neurons, in an attempt to insert it in a retroviral vector for use in overexpression experiments. In the RT-qPCR results showed before, increased levels of Tuj1 can be identified when applying Sox2 retroviral reprogramming. Tuj1 is a common marker for immature neurons, which indicates that HuFs are either being trans-differentiated to neuronal-committed cells or reprogrammed iNSCs are rapidly undergoing neuronal differentiation under culture conditions. Overexpressing Pnky, whose downregulation in NSCs leads do increased neuronal commitment could help counter this effect by repressing neuronal specific programs.

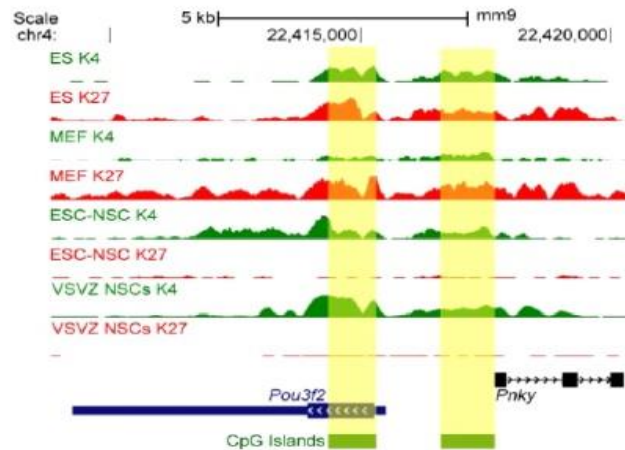


Figure 3.12 – H3K4me3 and H3K27me3 ChIP-seq tracks for embryonic stem cells (ES), mouse embryonic fibroblasts (MEF), ESC-derived neural stem cells (ESC-NSCs), and V-SVZ NSCs of the *Pnky* and *Pou3f2* loci viewed in Genome browser with promoter regions highlighted in yellow (Ramos et al. 2015).

Since *Pnky* is only expressed in the brain, we considered starting with a mouse model approach, so we cloned mouse *Pnky* from a RNA sample obtained from a C57BL/6 mouse dissected brain SVZ. Obtained cDNA from reverse transcription was further amplified by PCR and purified by electrophoresis, the obtained 825nt band is presented in **Figure 3.13 A**, and *Pnky* identity was confirmed by BLAST analysis of Sanger sequencing results.

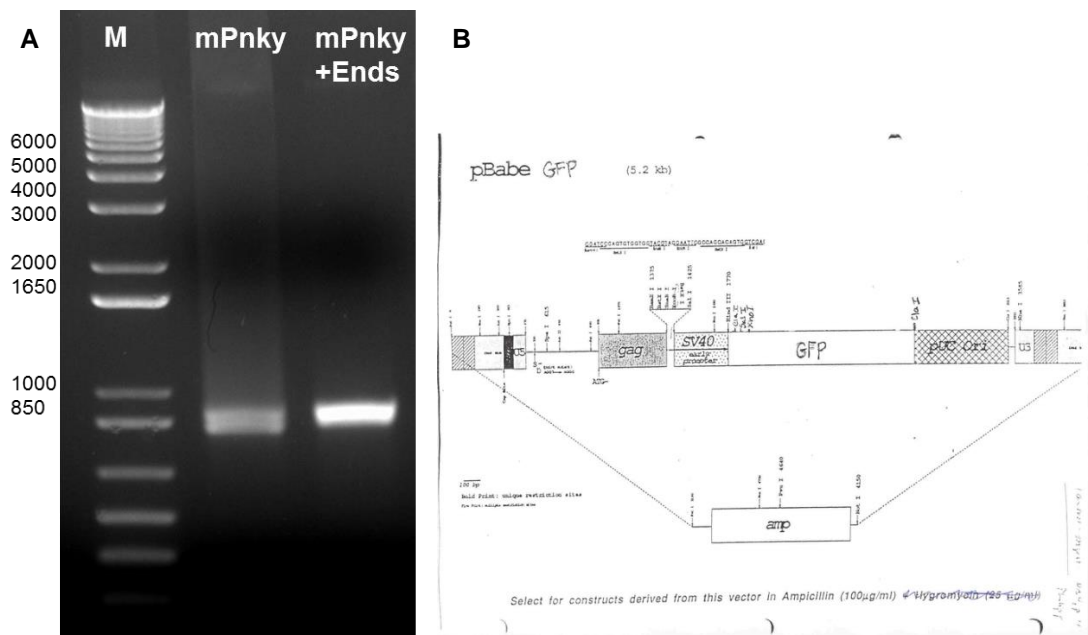


Figure 3.13 – **A** PCR product of SVZ cDNA amplified with mPnky homology primers and ligation ends with *Bst*XI restriction sites, purified and run on agarose 1% TAE gel (image taken with ChemiDoc™ XRS+ System Bio-Rad). **B** pBabeGFP plasmid design (addgene), plasmid was restricted with *Eco*RI and *Cla*I, excised GFP was substituted by mPnky.

Our first attempt to obtain a retroviral plasmid for overexpressing mouse *Pnky* (mPnky) was performed by ligation of mPnky amplified by PCR with primers that included a restriction site for *Bst*XI restriction enzyme, producing sticky ends for orientated ligation to pMXs vector plasmid. However, this

approach showed to be difficult as after various attempts to ligate plasmid and insert failed. Various vector:insert ratios were attempted from 1:1 to 1:10, but the insert never made it into the vector as accessed by diagnostic restriction and Sanger sequencing of obtained plasmids from DH5 α bacterial transformation, the reasons behind this outcome are still unclear.

Since this approach was inefficient, blunt end ligation is currently being tried, using another vector plasmid. The pBabeGFP plasmid was chosen as the alternative since it had been used before for this purpose. Ligations are still being tried, so far we were able to get the insert inside the plasmid but the insertion occurred in the wrong orientation, a disadvantage of blunt end ligation. Still, it should be a matter of time before a functional mPnky expression plasmid is obtained.

Recently, we also managed to start cloning the human Pnky from glioblastoma cells RNA sample, which if successful, should provide us with a more easily adaptable experience model, considering our already in progress experiments with human fibroblast reprogramming.

Future Prospects

As soon as a functional overexpression plasmid for Pnky is obtained, experiments will start with the presented protocol for iNSC reprogramming, this time with an overpression approach with Pnky, who seems like a more suitable candidate for this particular experiment, since as indicated by our preliminary RT-qPCR results, Tuj1 is one of the most expressed markers by our cells. Countering this apparent tendency could improve efficiency and increase stemness properties by blocking neuronal differentiation. The role of Zeb2NAT in this approach also needs to be further explored, as results so far are inconclusive of it's effect in neural reprogramming.

4. Conclusions

Direct conversion allows us to obtain multipotent progenitors of various cell lines bypassing the pluripotent state, which presents enormous potential for medical applications such as tissue recovery therapies. However most reprogramming protocols still struggle with low efficiencies, particularly when reprogramming older patient's cells. Since long noncoding RNAs are recently being related to diverse regulatory functions in cell character definition and differentiation, understanding how these molecules contribute to direct conversion could prove to be useful in finding ways to improve reprogramming protocols.

Reprogramming human fibroblasts directly to lineage committed progenitors using only one pluripotency transcription factor seems to involve specific expression dynamics. Although epigenetic activation can be performed by Oct4 or Sox2 individually, the cellular mechanisms involved in redirecting cells toward a different lineage aren't necessarily the same as when using transient expression of OSKM factors. When using the OSKM method, the early MET is a crucial step in reprogramming efficiency, however when using only one of the main pluripotency TFs Oct4 or Sox2 to promote epigenetic activation it is not clear if this transition is essential or even needed. Downregulating Zeb2NAT LncRNA during early reprogramming seems to affect colony formation in iHPC reprogramming with Oct4, however the challenge in combining LNA treatment with retroviral expression during iNSC reprogramming raises questions about the retroviral expression efficiency. This could be the reason behind the failure of our iHPC protocols first attempt, also Oct4 reprogramming could work by unspecific targeting of other POU domain targets, as Oct4 is not exactly a known intervenient of hematopoietic programs like Oct1 or Oct2. This could explain why using lentiviral vectors was much more efficient in the original protocol by Szabo *et al* when compared to our retroviral attempt since it would need higher Oct4 expression levels to activate unspecific pathways. In the future we hope to optimize the combined use of LNA GapmeRs and retroviral transduction, allowing us to determine definitely if the apparent interference with direct conversion protocols is due to LNA treatment or effectively a consequence of Zeb2NAT downregulation.

Although our results regarding iNSC reprogramming are unclear, it is very likely that Zeb2NAT expression affects reprogramming, the main question remains if downregulation is the right path to improve this direct conversion process, as the need for early MET is still questionable and Zeb2 expression is known to be needed in neural development during embryogenesis and in central nervous system early development. On this topic, overexpression could prove to be a more suitable approach, and LncRNA Pnky could be an interesting alternative in neural stem cell reprogramming. Overexpressing Pnky during reprogramming to NSCs using Sox2 could help getting higher yields by blocking trans-differentiation to neuronal fate, what seems to be an issue with this culture conditions as Tuj1 levels seem higher than Nestin's as accessed by RT-qPCR, although our housekeeping's are not completely reliable this goes in accordance with the changes observed in cell culture. If this approach is successful it would be interesting to characterize obtained NSCs and compare them to wild type SVZ-NSCs, as Pnky is highly expressed in this particular NSC niche.

Continuing to explore LncRNA potential as mediators of reprogramming can provide insight on the bigger picture of cell fate maintenance and regulation, opening new prospects in direct conversion. The work developed in this thesis is still in an early stage, some methods still need optimization but the means to answering the questions proposed and achieve better knowledge of LncRNA importance and contribution to direct conversion are within reach.

5. References

- Abad, M. et al., 2013. Reprogramming in vivo produces teratomas and iPS cells with totipotency features. *Nature*, 502(7471), pp.340–345. Available at: <http://www.nature.com/doi/10.1038/nature12586>.
- Ashapkin, V. V, Kutueva, L.I. & Vanyushin, B.F., 2015. Aging Epigenetics : Accumulation of Errors or Realization of a Specific Program ? , 80(11), pp.1406–1417.
- Beltran, M. et al., 2008. A natural antisense transcript regulates Zeb2 / Sip1 gene expression during Snail1-induced epithelial – mesenchymal transition. *Genes and Development*, 22, pp.756–769.
- Ben-David, U. & Benvenisty, N., 2011. The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nature Reviews Cancer*, 11(4), pp.268–277. Available at: <http://www.nature.com/doi/10.1038/nrc3034>.
- Goossens, S. et al., 2011. The EMT regulator Zeb2/Sip1 is essential for murine embryonic hematopoietic stem/progenitor cell differentiation and mobilization. *Blood*, 117(21), pp.5620–5630.
- Guttman, M. et al., 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*, 477(7364), pp.295–300. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3175327&tool=pmcentrez&rendertype=abstract>.
- Hegarty, S. V, Sullivan, A.M. & O’Keeffe, G.W., 2015. Zeb2: A multifunctional regulator of nervous system development. *Progress in neurobiology*, 132, pp.1–15. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26193487>.
- Johnsson, P. et al., 2014. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(3), pp.1063–1071. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0304416513004753>.
- Kauppinen, S., Vester, B. & Wengel, J., 2005. Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics. *Drug Discovery Today: Technologies*, 2(3), pp.279–282.
- Kawahara, H., Imai, T. & Okano, H., 2012. MicroRNAs in neural stem cells and neurogenesis. *Frontiers in Neuroscience*, 6(MAR), pp.1–13.
- Kelaini, S., Cochrane, A. & Margariti, A., 2014. Direct reprogramming of adult cells: Avoiding the pluripotent state. *Stem Cells and Cloning: Advances and Applications*, 7(1), pp.19–29.
- Li, G. et al., 2013. Multipotent Neural Stem Cells with a Single Factor. , 11(1), pp.100–109.
- De Los Angeles, A. et al., 2015. Hallmarks of pluripotency. *Nature*, 525(7570), pp.469–78. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26399828>.
- Ma, T. et al., 2013. Progress in the reprogramming of somatic cells. *Circulation Research*, 112(3), pp.562–574.
- Ma, Z.Z. et al., 2015. A novel method to derive and expand mice neural stem cells efficiently without neuro-sphere formation. *International Journal of Clinical and Experimental Medicine*, 8(8), pp.12834–12841.
- Matsui, T. et al., 2012. Neural Stem Cells Directly Differentiated from Partially Reprogrammed Fibroblasts Rapidly Acquire Gliogenic Competency. *Stem Cells*, 30(6), pp.1109–1119. Available at: <http://doi.wiley.com/10.1002/stem.1091>.
- Nichols, J. & Smith, A., 2009. Naive and Primed Pluripotent States. *Cell Stem Cell*, 4(6), pp.487–492. Available at: <http://dx.doi.org/10.1016/j.stem.2009.05.015>.
- Park, S. et al., 2008. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. , pp.894–907.
- Pauli, A., Rinn, J.L. & Schier, A.F., 2011. Non-coding RNAs as regulators of embryogenesis. *Nature reviews. Genetics*, 12(2), pp.136–49. Available at: <http://www.nature.com/pros.lib.unimi.it/nrg/journal/v12/n2/full/nrg2904.html>.
- Ramos, A.D. et al., 2015. The long noncoding RNA Pnky regulates neuronal differentiation of embryonic and postnatal neural stem cells. *Cell Stem Cell*, 16, pp.439–447. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25800779>.
- Ring, K.L. et al., 2012. Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. *Cell Stem Cell*, 11(1), pp.100–109. Available at:

- <http://dx.doi.org/10.1016/j.stem.2012.05.018>.
- Szabo, E. et al., 2010. Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature*, 468(7323), pp.521–526. Available at: <http://www.nature.com/doifinder/10.1038/nature09591>.
- Takahashi, K. & Yamanaka, S., 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4), pp.663–676.
- Trokovic, R. et al., 2015. Combined negative effect of donor age and time in culture on the reprogramming efficiency into induced pluripotent stem cells. *Stem cell research*, 15(1), pp.254–262. Available at: <http://www.sciencedirect.com/science/article/pii/S1873506115000744>.
- Vandewalle, C. et al., 2005. SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic acids research*, 33(20), pp.6566–78. Available at: <http://nar.oxfordjournals.org/content/33/20/6566.full>.
- Wang, G. et al., 2013. Critical regulation of miR-200 / ZEB2 pathway in Oct4 / Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(8), pp.2858–63.
- Wang, Y. et al., 2013. Endogenous miRNA Sponge lincRNA-RoR Regulates Oct4, Nanog, and Sox2 in Human Embryonic Stem Cell Self-Renewal. *Developmental Cell*, 25(1), pp.69–80. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1534580713001330>.
- Xue, Y. et al., 2013. Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated MicroRNA circuits. *Cell*, 152(1-2), pp.82–96.
- Yamanaka, S. et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5), pp.861–72. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18035408>.
- Yang, Y. et al., Enhanced rejuvenation in iPSC-derived neurons compared with directly converted neurons from an aged mouse. , pp.1–44. *Stem Cells Dev.* 2015 Dec 1;24(23):2767-77.

6. Annexes

Primer Sequences (5'-3')

mPnky Cloning primers (100% homology)

Forward: GGGAGAAGCAACTTCCTCTGG

Reverse: ACAATTAATTTTTTCAGGTAAATTTTATTT

mPnky insert primers (with BstXI restriction site)

Forward: ATACCACTGTGCTGGGGGAGAAGCAACT

Reverse: CTACCACCACACTGGACAATTAATTTT

pMXs plasmid sequencing primer:

pMX-L3205: CCCTTTTCTGGAGACTAAATAAA

pBabe sequencing primers:

Forward: CTTTATCCAGCCCTCAC

Reverse: ACCCTAACTGACACACATTCC

qPCR primer sequences (Invitrogen):

hNestin F – ACCAAGAGACATTCAGACTCC

hNestin R – CCTCATCCTCATTTTCCACTCC

hSox2 F – AGCTACAGCATGATGCAGGA

hSox2 R – GGTCATGGAGTTGTACTGCA

hOct4 F – CGTTCTCTTTGGAAAGGTGTTC

hOct4 R – ACACTCGGACCACGTCTTTC

hTuj1 F – GCCTCTTCTCACAAGTACGTGCCTCG

hTuj1 R – GGGGCGAAGCCGGGCATGAACAAGAAGTGCAG

hNurr1 F – CGACATTTCTGCCTTCTCC

hNurr1 R – GGTAAGGTGTCCAGGAAAAG

hPAX6 F – AACAGACACAGCCCTCACAAACA

hPAX6 R – CGGGAAGTTGAACTGGAAGTAC

hGAPDH F – GTGGACCTGACCTGCCGTCT

hGAPDH R – GGAGGAGTGGGTGTCTGCTGT

hActin F – CAAGGCCAACC GCGAGAAGAT

hActin R – CCAGAGGCGTACAGGGATAGCAC

hZeb2NAT F – ACAAAGATAGGTGGCGCGTG

hZeb2NAT R – GCATGAAGAAGCCGCGAAGTGT

hMAP2 F – CCACCTGAGATTAAGGATCA

hMAP2 R – GGCTTACTTTGCTTCTCTGA

hGFAP F – GTACCAGGACCTGCTCAAT

hGFAP R – CAACTATCCTGCTTCTGCTC

LNA GapmeR sequences (5'-3')

Scramble LNA: AACACGTCTATACGC

hZeb2NAT1: GTCCAGAAATTCATC

Zeb2NAT2: TTAGTGATGAGGATA

mPnky cDNA sequence

GGGAGAAGCA ACTTCCTCTG GTCTTCTGGA GGTGTAACCT ACGTGCCCAG
TAGGATATAC TGCCGGGTTG TGAAATGTCC ACGCCTCTCC CCAACTGTCT
TCCTTCCCAG CCCTCCGGGC TTGGCTTTCT TGCTTCCCCA GGAGTTCAAA
TCTCCAACCT GCGGAAGAAT TCAGCTGCTT GAAAGGACTT AAGGCAGTGT
GCGGAGGACA TCTCCTTTCT CCGCCAGTAA AGAGAGCTGT TCAAAGACCG
AGGCTGCTTA CGATGACGTG GAGAGGATTT CAAACAACCT TAACAGGCTT
TGAAGTACACA AGAAGCGGTG ATATCTGCCT CACTCTGGCA CAGCTCCTCC
AGTGCACTTG CTAGGACAAT GGCTGAGAAA GCACTTGGTG CTGGCCTCTC
TGCCGGGGGC GAGGACTGCC CAAACCCAG AGACCCTAAG GCAGGAGTTG
CTGCACTACA ATGGAGAGGT CTTGTCTTAG GCTGCACCTC AGGTTCTCTC
AGATCTTCTG GAAGGGCGTT ATTCAGCGGT ACTGTGGTTG CGCTGTGCCA
GCAGCTGCTT GATGGAGACC TCTTCTCCCC ACATCTGAAT GGAACGTCTT
TGCCCAGAGT CTACAGAATG TCAAACTGA GGCTCCGACC TCAGAGCTAC
AGCTTTAAGG ACCACTCCAC ACAAAGAGGC AGCTGGTTGC TTTAATGAAA
ACTGCCTTTA AGCTTCAAGA ACTGAGGCCT TGGGGAATCC ATTTATAAGG
AGCCTAGAAA ATGCATTTCC AAGTTGTATG TTCTTAGGAG ATACAAAATA
AAATTTACCT GAAAAAATTA ATTGT

iHPCs

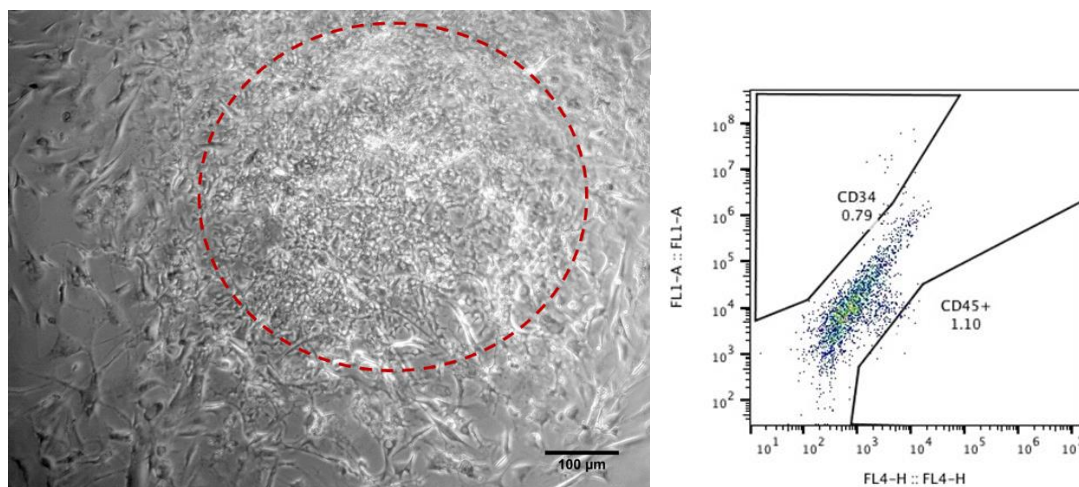


Figure 6.1 - Colony forming from Oct4 transduced 3YR HuFs after 20 days in differentiation culture conditions and FACS analysis profile of same cells stained with anti-humanCD45 and anti-humanCD34 antibodies.

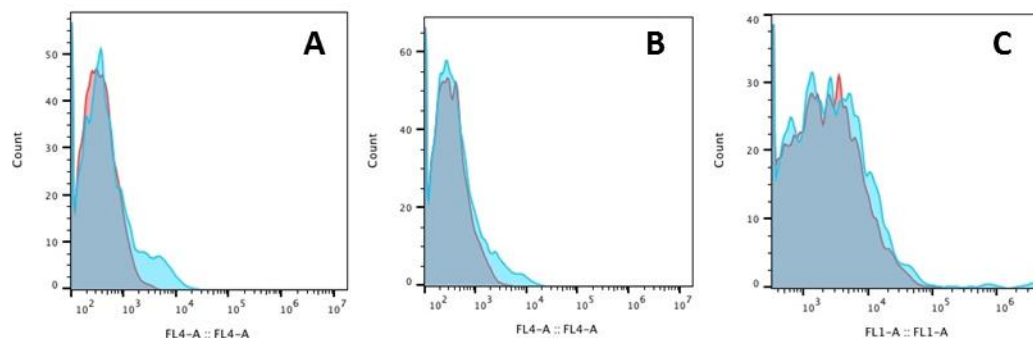


Figure 6.2 - Histogram representation of 3YR cells reprogramed with Oct4 (Red for unstained and Blue for cells stained with APC-antiCD45 antibody). **A** Oct4 alone, **B** Oct4 + LNA targeting Zeb2NAT and **C** Oct4 + Scramble LNA.

iNSCs

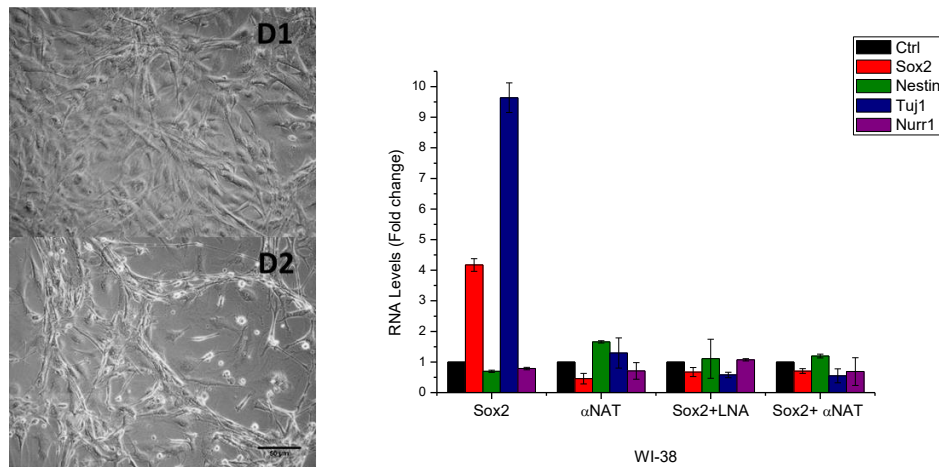


Figure 6.3 - Control WI-38 HuFs transduced with GFP retrovirus alone after 3 days (D1) and 9days (D2) in culture with neural differentiation medium and RT-qPCR result expression levels, presented as $\Delta\Delta C_t$ using Actin as housekeeping gene.

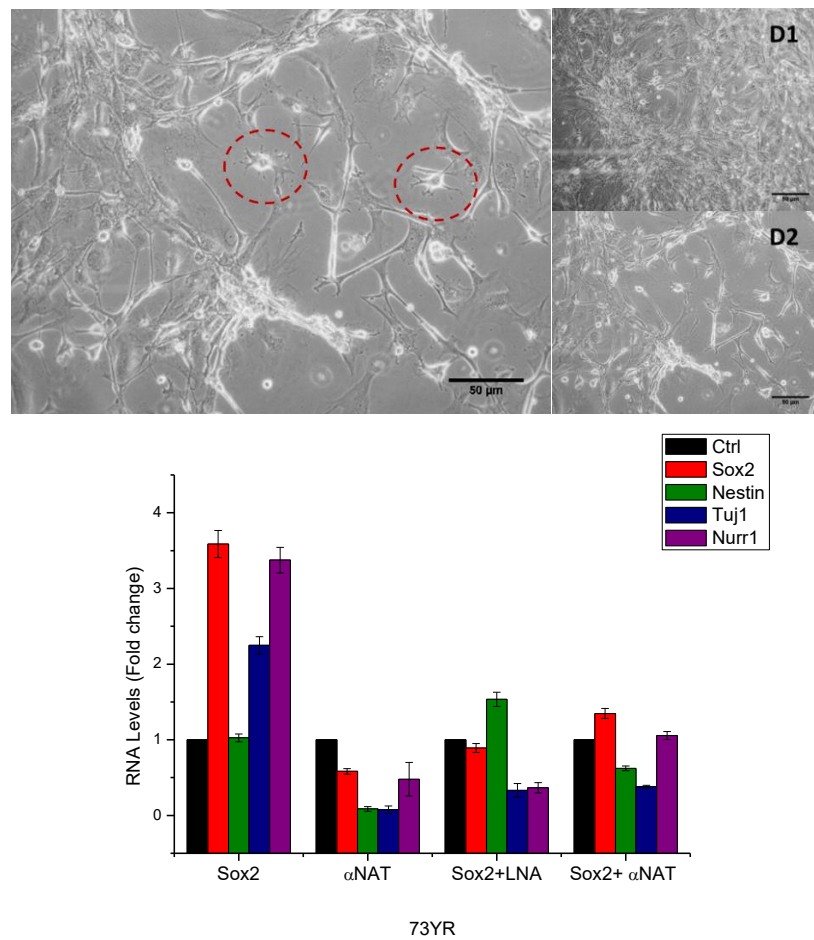


Figure 6.4 - Control 73YR HuFs transduced with GFP retrovirus alone after 3 days (D1) and 9days (D2) in culture with neural differentiation medium and RT-qPCR result expression levels, presented as $\Delta\Delta C_t$ using Actin as housekeeping gene.